

1269017

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME;

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

January 04, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/524,587

FILING DATE: *November 24, 2003*

RELATED PCT APPLICATION NUMBER: *PCT/US04/39549*



Certified By

Jon W Dudas

Under Secretary
of Commerce for Intellectual Property
and Acting Director of the
United States Patent and Trademark Office



15750 U.S. PTO

PTO/SB/16 (08-03)

Approved for use through 07/31/2006. OMB 0551-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EV342494605US

60/524587



112403

INVENTOR(S)					
Given Name (first and middle (if any))		Family Name or Surname		Residence (City and either State or Foreign Country)	
Michael R.		Costa		San Francisco, California	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto					
TITLE OF THE INVENTION (500 characters max)					
MBCATs AS MODIFIERS OF THE BETA-CATENIN PATHWAY AND METHODS OF USE					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number: <div style="border: 1px solid black; padding: 5px; display: inline-block;">23500</div>					
OR					
<input type="checkbox"/> Firm or Individual Name					
Address					
Address					
City		State		Zip	
Country		Telephone		Fax	
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification Number of Pages <u>73</u>		<input type="checkbox"/> CD(s), Number _____			
<input type="checkbox"/> Drawing(s) Number of Sheets _____		<input checked="" type="checkbox"/> Other (specify) <u>return postcard</u>			
<input type="checkbox"/> Application Date Sheet. See 37 CFR 1.76					
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT					
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE Amount (\$)	
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees.				<div style="border: 1px solid black; padding: 10px; text-align: center;">160.00</div>	
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: <u>50-1108</u>					
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.					
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/> No.					
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____					

[Page 1 of 2]

Respectfully submitted,

SIGNATURE

TYPED or PRINTED NAME Laleh Shayesteh

TELEPHONE 650.837.8223

Date November 24, 2003

REGISTRATION NO. 47,937

(if appropriate)

Docket Number: EX03-092P

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/16 (08-03)

Approved for use through 07/31/2006. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number **EX03-092P**

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle (if any))	Family or Surname	Residence (City and either State or Foreign Country)
Steven Brian	Gendreau	San Francisco, California
Emery G. III	Dora	San Francisco, California
HaiGuang	Zhang	El Sobrante, California
Lynn Margaret	Bjerke	San Francisco, California
Joanne I.	Adamkewicz	San Francisco, California
Kim	Lickteig	San Francisco, California

[Page 2 of 2]

Number 1 of 1

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

MBCATs AS MODIFIERS OF THE BETA-CATENIN PATHWAY AND METHODS OF USE

BACKGROUND OF THE INVENTION

The *Drosophila Melanogaster* Armadillo/beta-catenin protein is implicated in multiple cellular functions. The protein functions in cell signaling via the Wingless (Wg)/Wnt signaling pathway. It also functions as a cell adhesion protein at the cell membrane in a complex with E-cadherin and alpha-catenin (Cox et al. (1996) J. Cell Biol. 134: 133-148; Godt and Tepass (1998) Nature 395: 387-391; White et al. (1998) J Cell Biol. 140:183-195). These two roles of beta -catenin can be separated from each other (Orsulic and Peifer (1996) J. Cell Biol. 134: 1283-1300; Sanson et al. (1996) Nature 383: 627-630).

In Wingless cell signaling, beta -catenin levels are tightly regulated by a complex containing APC, Axin, and GSK3 beta /SGG/ZW3 (Peifer et al. (1994) Development 120: 369-380).

The Wingless/ beta -catenin signaling pathway is frequently mutated in human cancers, particularly those of the colon. Mutations in the tumor suppressor gene APC, as well as point mutations in beta -catenin itself lead to the stabilization of the beta -catenin protein and inappropriate activation of this pathway.

The ability to manipulate the genomes of model organisms such as *Drosophila* provides a powerful means to analyze biochemical processes that, due to significant evolutionary conservation, have direct relevance to more complex vertebrate organisms. Due to a high level of gene and pathway conservation, the strong similarity of cellular processes, and the functional conservation of genes between these model organisms and mammals, identification of the involvement of novel genes in particular pathways and their functions in such model organisms can directly contribute to the understanding of the correlative pathways and methods of modulating them in mammals (see, for example, Mechler BM et al., 1985 EMBO J 4:1551-1557; Gateff E. 1982 Adv. Cancer Res. 37: 33-74; Watson KL., et al., 1994 J Cell Sci. 18: 19-33; Miklos GL, and Rubin GM. 1996 Cell 86:521-529; Wassarman DA, et al., 1995 Curr Opin Gen Dev 5: 44-50; and Booth DR. 1999 Cancer

Metastasis Rev. 18: 261-284). For example, a genetic screen can be carried out in an invertebrate model organism having underexpression (e.g. knockout) or overexpression of a gene (referred to as a “genetic entry point”) that yields a visible phenotype. Additional genes are mutated in a random or targeted manner. When a gene mutation changes the original phenotype caused by the mutation in the genetic entry point, the gene is identified as a “modifier” involved in the same or overlapping pathway as the genetic entry point. When the genetic entry point is an ortholog of a human gene implicated in a disease pathway, such as beta-catenin, modifier genes can be identified that may be attractive candidate targets for novel therapeutics.

All references cited herein, including patents, patent applications, publications, and sequence information in referenced Genbank identifier numbers, are incorporated herein in their entireties.

SUMMARY OF THE INVENTION

We have discovered genes that modify the beta-catenin pathway in *Drosophila*, and identified their human orthologs, hereinafter referred to as modifier of beta catenin (MBCAT). The invention provides methods for utilizing these beta-catenin modifier genes and polypeptides to identify MBCAT-modulating agents that are candidate therapeutic agents that can be used in the treatment of disorders associated with defective or impaired beta-catenin function and/or MBCAT function. Preferred MBCAT-modulating agents specifically bind to MBCAT polypeptides and restore beta-catenin function. Other preferred MBCAT-modulating agents are nucleic acid modulators such as antisense oligomers and RNAi that repress MBCAT gene expression or product activity by, for example, binding to and inhibiting the respective nucleic acid (i.e. DNA or mRNA).

MBCAT modulating agents may be evaluated by any convenient *in vitro* or *in vivo* assay for molecular interaction with an MBCAT polypeptide or nucleic acid. In one embodiment, candidate MBCAT modulating agents are tested with an assay system comprising a MBCAT polypeptide or nucleic acid. Agents that produce a change in the activity of the assay system relative to controls are identified as candidate beta-catenin modulating agents. The assay system may be cell-based or

cell-free. MBCAT-modulating agents include MBCAT related proteins (e.g. dominant negative mutants, and biotherapeutics); MBCAT -specific antibodies; MBCAT -specific antisense oligomers and other nucleic acid modulators; and chemical agents that specifically bind to or interact with MBCAT or compete with MBCAT binding partner (e.g. by binding to an MBCAT binding partner). In one specific embodiment, a small molecule modulator is identified using a binding assay. In specific embodiments, the screening assay system is selected from an apoptosis assay, a cell proliferation assay, an angiogenesis assay, and a hypoxic induction assay.

In another embodiment, candidate beta-catenin pathway modulating agents are further tested using a second assay system that detects changes in the beta-catenin pathway, such as angiogenic, apoptotic, or cell proliferation changes produced by the originally identified candidate agent or an agent derived from the original agent. The second assay system may use cultured cells or non-human animals. In specific embodiments, the secondary assay system uses non-human animals, including animals predetermined to have a disease or disorder implicating the beta-catenin pathway, such as an angiogenic, apoptotic, or cell proliferation disorder (e.g. cancer).

The invention further provides methods for modulating the MBCAT function and/or the beta-catenin pathway in a mammalian cell by contacting the mammalian cell with an agent that specifically binds a MBCAT polypeptide or nucleic acid. The agent may be a small molecule modulator, a nucleic acid modulator, or an antibody and may be administered to a mammalian animal predetermined to have a pathology associated with the beta-catenin pathway.

DETAILED DESCRIPTION OF THE INVENTION

In a screen to identify enhancers and suppressors of the Wg signaling pathway, we generated activated beta -catenin models in *Drosophila* based on human tumor data (Polakis (2000) Genes and Development 14: 1837-1851). We identified modifiers of the Wg pathway and identified their orthologs. Accordingly, vertebrate orthologs of these modifiers, and preferably the human orthologs, MBCAT genes (i.e., nucleic acids and polypeptides) are attractive drug targets for the treatment of

pathologies associated with a defective beta-catenin signaling pathway, such as cancer. Table 1 (Example II) lists the modifiers and their orthologs.

In vitro and in vivo methods of assessing MBCAT function are provided herein. Modulation of the MBCAT or their respective binding partners is useful for understanding the association of the beta-catenin pathway and its members in normal and disease conditions and for developing diagnostics and therapeutic modalities for beta-catenin related pathologies. MBCAT-modulating agents that act by inhibiting or enhancing MBCAT expression, directly or indirectly, for example, by affecting an MBCAT function such as enzymatic (e.g., catalytic) or binding activity, can be identified using methods provided herein. MBCAT modulating agents are useful in diagnosis, therapy and pharmaceutical development.

Nucleic acids and polypeptides of the invention

Sequences related to MBCAT nucleic acids and polypeptides that can be used in the invention are disclosed in Genbank (referenced by Genbank identifier (GI) or RefSeq number), shown in Table 1 and in the appended sequence listing.

The term “MBCAT polypeptide” refers to a full-length MBCAT protein or a functionally active fragment or derivative thereof. A “functionally active” MBCAT fragment or derivative exhibits one or more functional activities associated with a full-length, wild-type MBCAT protein, such as antigenic or immunogenic activity, enzymatic activity, ability to bind natural cellular substrates, etc. The functional activity of MBCAT proteins, derivatives and fragments can be assayed by various methods known to one skilled in the art (Current Protocols in Protein Science (1998) Coligan *et al.*, eds., John Wiley & Sons, Inc., Somerset, New Jersey) and as further discussed below. In one embodiment, a functionally active MBCAT polypeptide is an MBCAT derivative capable of rescuing defective endogenous MBCAT activity, such as in cell based or animal assays; the rescuing derivative may be from the same or a different species. For purposes herein, functionally active fragments also include those fragments that comprise one or more structural domains of an MBCAT, such as a kinase domain or a binding domain. Protein domains can be identified using the PFAM program (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2). Methods

for obtaining MBCAT polypeptides are also further described below. In some embodiments, preferred fragments are functionally active, domain-containing fragments comprising at least 25 contiguous amino acids, preferably at least 50, more preferably 75, and most preferably at least 100 contiguous amino acids of an MBCAT. In further preferred embodiments, the fragment comprises the functionally active domain.

The term "MBCAT nucleic acid" refers to a DNA or RNA molecule that encodes a MBCAT polypeptide. Preferably, the MBCAT polypeptide or nucleic acid or fragment thereof is from a human, but can also be an ortholog, or derivative thereof with at least 70% sequence identity, preferably at least 80%, more preferably 85%, still more preferably 90%, and most preferably at least 95% sequence identity with human MBCAT. Methods of identifying orthologs are known in the art. Normally, orthologs in different species retain the same function, due to presence of one or more protein motifs and/or 3-dimensional structures. Orthologs are generally identified by sequence homology analysis, such as BLAST analysis, usually using protein bait sequences. Sequences are assigned as a potential ortholog if the best hit sequence from the forward BLAST result retrieves the original query sequence in the reverse BLAST (Huynen MA and Bork P, *Proc Natl Acad Sci* (1998) 95:5849-5856; Huynen MA *et al.*, *Genome Research* (2000) 10:1204-1210). Programs for multiple sequence alignment, such as CLUSTAL (Thompson JD *et al.*, 1994, *Nucleic Acids Res* 22:4673-4680) may be used to highlight conserved regions and/or residues of orthologous proteins and to generate phylogenetic trees. In a phylogenetic tree representing multiple homologous sequences from diverse species (e.g., retrieved through BLAST analysis), orthologous sequences from two species generally appear closest on the tree with respect to all other sequences from these two species. Structural threading or other analysis of protein folding (e.g., using software by ProCeryon, Biosciences, Salzburg, Austria) may also identify potential orthologs. In evolution, when a gene duplication event follows speciation, a single gene in one species, such as *Drosophila*, may correspond to multiple genes (paralogs) in another, such as human. As used herein, the term "orthologs" encompasses paralogs. As used herein, "percent (%) sequence identity" with respect to a subject sequence, or a

specified portion of a subject sequence, is defined as the percentage of nucleotides or amino acids in the candidate derivative sequence identical with the nucleotides or amino acids in the subject sequence (or specified portion thereof), after aligning the sequences and introducing gaps, if necessary to achieve the maximum percent sequence identity, as generated by the program WU-BLAST-2.0a19 (Altschul *et al.*, J. Mol. Biol. (1997) 215:403-410) with all the search parameters set to default values. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. A % identity value is determined by the number of matching identical nucleotides or amino acids divided by the sequence length for which the percent identity is being reported. "Percent (%) amino acid sequence similarity" is determined by doing the same calculation as for determining % amino acid sequence identity, but including conservative amino acid substitutions in addition to identical amino acids in the computation.

A conservative amino acid substitution is one in which an amino acid is substituted for another amino acid having similar properties such that the folding or activity of the protein is not significantly affected. Aromatic amino acids that can be substituted for each other are phenylalanine, tryptophan, and tyrosine; interchangeable hydrophobic amino acids are leucine, isoleucine, methionine, and valine; interchangeable polar amino acids are glutamine and asparagine; interchangeable basic amino acids are arginine, lysine and histidine; interchangeable acidic amino acids are aspartic acid and glutamic acid; and interchangeable small amino acids are alanine, serine, threonine, cysteine and glycine.

Alternatively, an alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman (Smith and Waterman, 1981, *Advances in Applied Mathematics* 2:482-489; database: European Bioinformatics Institute; Smith and Waterman, 1981, *J. of Molec.Biol.*, 147:195-197; Nicholas et al., 1998, "A Tutorial on Searching Sequence Databases and Sequence Scoring Methods" (www.psc.edu) and references cited therein.; W.R. Pearson, 1991, *Genomics* 11:635-650). This algorithm can be applied to amino acid sequences by using the scoring

matrix developed by Dayhoff (Dayhoff: Atlas of Protein Sequences and Structure, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA), and normalized by Gribskov (Gribskov 1986 Nucl. Acids Res. 14(6):6745-6763). The Smith-Waterman algorithm may be employed where default parameters are used for scoring (for example, gap open penalty of 12, gap extension penalty of two). From the data generated, the "Match" value reflects "sequence identity."

Derivative nucleic acid molecules of the subject nucleic acid molecules include sequences that hybridize to the nucleic acid sequence of an MBCAT. The stringency of hybridization can be controlled by temperature, ionic strength, pH, and the presence of denaturing agents such as formamide during hybridization and washing. Conditions routinely used are set out in readily available procedure texts (*e.g.*, Current Protocol in Molecular Biology, Vol. 1, Chap. 2.10, John Wiley & Sons, Publishers (1994); Sambrook *et al.*, Molecular Cloning, Cold Spring Harbor (1989)). In some embodiments, a nucleic acid molecule of the invention is capable of hybridizing to a nucleic acid molecule containing the nucleotide sequence of an MBCAT under high stringency hybridization conditions that are: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C in a solution comprising 6X single strength citrate (SSC) (1X SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5X Denhardt's solution, 0.05% sodium pyrophosphate and 100 µg/ml herring sperm DNA; hybridization for 18-20 hours at 65° C in a solution containing 6X SSC, 1X Denhardt's solution, 100 µg/ml yeast tRNA and 0.05% sodium pyrophosphate; and washing of filters at 65° C for 1h in a solution containing 0.1X SSC and 0.1% SDS (sodium dodecyl sulfate).

In other embodiments, moderately stringent hybridization conditions are used that are: pretreatment of filters containing nucleic acid for 6 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH7.5), 5mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 µg/ml denatured salmon sperm DNA; hybridization for 18-20h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH7.5), 5mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100

$\mu\text{g/ml}$ salmon sperm DNA, and 10% (wt/vol) dextran sulfate; followed by washing twice for 1 hour at 55° C in a solution containing 2X SSC and 0.1% SDS.

Alternatively, low stringency conditions can be used that are: incubation for 8 hours to overnight at 37° C in a solution comprising 20% formamide, 5 x SSC, 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 $\mu\text{g/ml}$ denatured sheared salmon sperm DNA; hybridization in the same buffer for 18 to 20 hours; and washing of filters in 1 x SSC at about 37° C for 1 hour.

Isolation, Production, Expression, and Mis-expression of MBCAT

Nucleic Acids and Polypeptides

MBCAT nucleic acids and polypeptides are useful for identifying and testing agents that modulate MBCAT function and for other applications related to the involvement of MBCAT in the beta-catenin pathway. MBCAT nucleic acids and derivatives and orthologs thereof may be obtained using any available method. For instance, techniques for isolating cDNA or genomic DNA sequences of interest by screening DNA libraries or by using polymerase chain reaction (PCR) are well known in the art. In general, the particular use for the protein will dictate the particulars of expression, production, and purification methods. For instance, production of proteins for use in screening for modulating agents may require methods that preserve specific biological activities of these proteins, whereas production of proteins for antibody generation may require structural integrity of particular epitopes.

Expression of proteins to be purified for screening or antibody production may require the addition of specific tags (*e.g.*, generation of fusion proteins).

Overexpression of an MBCAT protein for assays used to assess MBCAT function, such as involvement in cell cycle regulation or hypoxic response, may require expression in eukaryotic cell lines capable of these cellular activities. Techniques for the expression, production, and purification of proteins are well known in the art; any suitable means therefore may be used (*e.g.*, Higgins SJ and Hames BD (eds.) *Protein Expression: A Practical Approach*, Oxford University Press Inc., New York 1999; Stanbury PF et al., *Principles of Fermentation Technology*, 2nd edition, Elsevier Science, New York, 1995; Doonan S (ed.) *Protein Purification Protocols*, Humana

Press, New Jersey, 1996; Coligan JE et al, Current Protocols in Protein Science (eds.), 1999, John Wiley & Sons, New York). In particular embodiments, recombinant MBCAT is expressed in a cell line known to have defective beta-catenin function. The recombinant cells are used in cell-based screening assay systems of the invention, as described further below.

The nucleotide sequence encoding an MBCAT polypeptide can be inserted into any appropriate expression vector. The necessary transcriptional and translational signals, including promoter/enhancer element, can derive from the native MBCAT gene and/or its flanking regions or can be heterologous. A variety of host-vector expression systems may be utilized, such as mammalian cell systems infected with virus (*e.g.* vaccinia virus, adenovirus, *etc.*); insect cell systems infected with virus (*e.g.* baculovirus); microorganisms such as yeast containing yeast vectors, or bacteria transformed with bacteriophage, plasmid, or cosmid DNA. An isolated host cell strain that modulates the expression of, modifies, and/or specifically processes the gene product may be used.

To detect expression of the MBCAT gene product, the expression vector can comprise a promoter operably linked to an MBCAT gene nucleic acid, one or more origins of replication, and, one or more selectable markers (*e.g.* thymidine kinase activity, resistance to antibiotics, *etc.*). Alternatively, recombinant expression vectors can be identified by assaying for the expression of the MBCAT gene product based on the physical or functional properties of the MBCAT protein in *in vitro* assay systems (*e.g.* immunoassays).

The MBCAT protein, fragment, or derivative may be optionally expressed as a fusion, or chimeric protein product (*i.e.* it is joined via a peptide bond to a heterologous protein sequence of a different protein), for example to facilitate purification or detection. A chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other using standard methods and expressing the chimeric product. A chimeric product may also be made by protein synthetic techniques, *e.g.* by use of a peptide synthesizer (Hunkapiller *et al.*, Nature (1984) 310:105-111).

Once a recombinant cell that expresses the MBCAT gene sequence is identified, the gene product can be isolated and purified using standard methods (*e.g.* ion exchange, affinity, and gel exclusion chromatography; centrifugation; differential solubility; electrophoresis). Alternatively, native MBCAT proteins can be purified from natural sources, by standard methods (*e.g.* immunoaffinity purification). Once a protein is obtained, it may be quantified and its activity measured by appropriate methods, such as immunoassay, bioassay, or other measurements of physical properties, such as crystallography.

The methods of this invention may also use cells that have been engineered for altered expression (mis-expression) of MBCAT or other genes associated with the beta-catenin pathway. As used herein, mis-expression encompasses ectopic expression, over-expression, under-expression, and non-expression (*e.g.* by gene knock-out or blocking expression that would otherwise normally occur).

Genetically modified animals

Animal models that have been genetically modified to alter MBCAT expression may be used in *in vivo* assays to test for activity of a candidate beta-catenin modulating agent, or to further assess the role of MBCAT in a beta-catenin pathway process such as apoptosis or cell proliferation. Preferably, the altered MBCAT expression results in a detectable phenotype, such as decreased or increased levels of cell proliferation, angiogenesis, or apoptosis compared to control animals having normal MBCAT expression. The genetically modified animal may additionally have altered beta-catenin expression (*e.g.* beta-catenin knockout). Preferred genetically modified animals are mammals such as primates, rodents (preferably mice or rats), among others. Preferred non-mammalian species include zebrafish, *C. elegans*, and *Drosophila*. Preferred genetically modified animals are transgenic animals having a heterologous nucleic acid sequence present as an extrachromosomal element in a portion of its cells, *i.e.* mosaic animals (see, for example, techniques described by Jakobovits, 1994, *Curr. Biol.* 4:761-763.) or stably integrated into its germ line DNA (*i.e.*, in the genomic sequence of most or all of its cells). Heterologous nucleic acid is

introduced into the germ line of such transgenic animals by genetic manipulation of, for example, embryos or embryonic stem cells of the host animal.

Methods of making transgenic animals are well-known in the art (for transgenic mice see Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442 (1985), U.S. Pat. Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Pat. No. 4,873,191 by Wagner et al., and Hogan, B., Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (1986); for particle bombardment see U.S. Pat. No., 4,945,050, by Sandford *et al.*; for transgenic *Drosophila* see Rubin and Spradling, Science (1982) 218:348-53 and U.S. Pat. No. 4,670,388; for transgenic insects see Berghammer A.J. *et al.*, A Universal Marker for Transgenic Insects (1999) Nature 402:370-371; for transgenic Zebrafish see Lin S., Transgenic Zebrafish, Methods Mol Biol. (2000);136:375-3830); for microinjection procedures for fish, amphibian eggs and birds see Houdebine and Chourrout, Experientia (1991) 47:897-905; for transgenic rats see Hammer *et al.*, Cell (1990) 63:1099-1112; and for culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection see, e.g., Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E. J. Robertson, ed., IRL Press (1987)). Clones of the nonhuman transgenic animals can be produced according to available methods (see Wilmut, I. *et al.* (1997) Nature 385:810-813; and PCT International Publication Nos. WO 97/07668 and WO 97/07669).

In one embodiment, the transgenic animal is a “knock-out” animal having a heterozygous or homozygous alteration in the sequence of an endogenous MBCAT gene that results in a decrease of MBCAT function, preferably such that MBCAT expression is undetectable or insignificant. Knock-out animals are typically generated by homologous recombination with a vector comprising a transgene having at least a portion of the gene to be knocked out. Typically a deletion, addition or substitution has been introduced into the transgene to functionally disrupt it. The transgene can be a human gene (e.g., from a human genomic clone) but more preferably is an ortholog of the human gene derived from the transgenic host species. For example, a mouse MBCAT gene is used to construct a homologous

recombination vector suitable for altering an endogenous MBCAT gene in the mouse genome. Detailed methodologies for homologous recombination in mice are available (see Capecchi, *Science* (1989) 244:1288-1292; Joyner *et al.*, *Nature* (1989) 338:153-156). Procedures for the production of non-rodent transgenic mammals and other animals are also available (Houdebine and Chourrout, *supra*; Pursel *et al.*, *Science* (1989) 244:1281-1288; Simms *et al.*, *Bio/Technology* (1988) 6:179-183). In a preferred embodiment, knock-out animals, such as mice harboring a knockout of a specific gene, may be used to produce antibodies against the human counterpart of the gene that has been knocked out (Claesson MH *et al.*, (1994) *Scan J Immunol* 40:257-264; Declerck PJ *et al.*, (1995) *J Biol Chem.* 270:8397-400).

In another embodiment, the transgenic animal is a "knock-in" animal having an alteration in its genome that results in altered expression (e.g., increased (including ectopic) or decreased expression) of the MBCAT gene, e.g., by introduction of additional copies of MBCAT, or by operatively inserting a regulatory sequence that provides for altered expression of an endogenous copy of the MBCAT gene. Such regulatory sequences include inducible, tissue-specific, and constitutive promoters and enhancer elements. The knock-in can be homozygous or heterozygous.

Transgenic nonhuman animals can also be produced that contain selected systems allowing for regulated expression of the transgene. One example of such a system that may be produced is the cre/loxP recombinase system of bacteriophage P1 (Lakso *et al.*, *PNAS* (1992) 89:6232-6236; U.S. Pat. No. 4,959,317). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355; U.S. Pat. No. 5,654,182). In a preferred embodiment, both Cre-LoxP and Flp-Frt are used in the same system to regulate expression of the transgene, and for sequential deletion of vector sequences in the same cell (Sun X *et al* (2000) *Nat Genet* 25:83-6).

The genetically modified animals can be used in genetic studies to further elucidate the beta-catenin pathway, as animal models of disease and disorders implicating defective beta-catenin function, and for *in vivo* testing of candidate therapeutic agents, such as those identified in screens described below. The candidate therapeutic agents are administered to a genetically modified animal having altered MBCAT function and phenotypic changes are compared with appropriate control animals such as genetically modified animals that receive placebo treatment, and/or animals with unaltered MBCAT expression that receive candidate therapeutic agent.

In addition to the above-described genetically modified animals having altered MBCAT function, animal models having defective beta-catenin function (and otherwise normal MBCAT function), can be used in the methods of the present invention. For example, a beta-catenin knockout mouse can be used to assess, *in vivo*, the activity of a candidate beta-catenin modulating agent identified in one of the *in vitro* assays described below. Preferably, the candidate beta-catenin modulating agent when administered to a model system with cells defective in beta-catenin function, produces a detectable phenotypic change in the model system indicating that the beta-catenin function is restored, i.e., the cells exhibit normal cell cycle progression.

Modulating Agents

The invention provides methods to identify agents that interact with and/or modulate the function of MBCAT and/or the beta-catenin pathway. Modulating agents identified by the methods are also part of the invention. Such agents are useful in a variety of diagnostic and therapeutic applications associated with the beta-catenin pathway, as well as in further analysis of the MBCAT protein and its contribution to the beta-catenin pathway. Accordingly, the invention also provides methods for modulating the beta-catenin pathway comprising the step of specifically modulating MBCAT activity by administering an MBCAT-interacting or -modulating agent.

As used herein, an "MBCAT-modulating agent" is any agent that modulates MBCAT function, for example, an agent that interacts with MBCAT to inhibit or enhance MBCAT activity or otherwise affect normal MBCAT function. MBCAT function can be affected at any level, including transcription, protein expression, protein localization, and

cellular or extra-cellular activity. In a preferred embodiment, the MBCAT - modulating agent specifically modulates the function of the MBCAT. The phrases "specific modulating agent", "specifically modulates", etc., are used herein to refer to modulating agents that directly bind to the MBCAT polypeptide or nucleic acid, and preferably inhibit, enhance, or otherwise alter, the function of the MBCAT. These phrases also encompass modulating agents that alter the interaction of the MBCAT with a binding partner, substrate, or cofactor (e.g. by binding to a binding partner of an MBCAT, or to a protein/binding partner complex, and altering MBCAT function). In a further preferred embodiment, the MBCAT- modulating agent is a modulator of the beta-catenin pathway (e.g. it restores and/or upregulates beta-catenin function) and thus is also a beta-catenin-modulating agent.

Preferred MBCAT-modulating agents include small molecule compounds; MBCAT-interacting proteins, including antibodies and other biotherapeutics; and nucleic acid modulators such as antisense and RNA inhibitors. The modulating agents may be formulated in pharmaceutical compositions, for example, as compositions that may comprise other active ingredients, as in combination therapy, and/or suitable carriers or excipients. Techniques for formulation and administration of the compounds may be found in "Remington's Pharmaceutical Sciences" Mack Publishing Co., Easton, PA, 19th edition.

Small molecule modulators

Small molecules are often preferred to modulate function of proteins with enzymatic function, and/or containing protein interaction domains. Chemical agents, referred to in the art as "small molecule" compounds are typically organic, non-peptide molecules, having a molecular weight up to 10,000, preferably up to 5,000, more preferably up to 1,000, and most preferably up to 500 daltons. This class of modulators includes chemically synthesized molecules, for instance, compounds from combinatorial chemical libraries. Synthetic compounds may be rationally designed or identified based on known or inferred properties of the MBCAT protein or may be identified by screening compound libraries. Alternative appropriate modulators of this class are natural products, particularly secondary metabolites from organisms such as plants or fungi, which can

also be identified by screening compound libraries for MBCAT-modulating activity. Methods for generating and obtaining compounds are well known in the art (Schreiber SL, Science (2000) 151: 1964-1969; Radmann J and Gunther J, Science (2000) 151:1947-1948).

Small molecule modulators identified from screening assays, as described below, can be used as lead compounds from which candidate clinical compounds may be designed, optimized, and synthesized. Such clinical compounds may have utility in treating pathologies associated with the beta-catenin pathway. The activity of candidate small molecule modulating agents may be improved several-fold through iterative secondary functional validation, as further described below, structure determination, and candidate modulator modification and testing. Additionally, candidate clinical compounds are generated with specific regard to clinical and pharmacological properties. For example, the reagents may be derivatized and re-screened using *in vitro* and *in vivo* assays to optimize activity and minimize toxicity for pharmaceutical development.

Protein Modulators

Specific MBCAT-interacting proteins are useful in a variety of diagnostic and therapeutic applications related to the beta-catenin pathway and related disorders, as well as in validation assays for other MBCAT-modulating agents. In a preferred embodiment, MBCAT-interacting proteins affect normal MBCAT function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In another embodiment, MBCAT-interacting proteins are useful in detecting and providing information about the function of MBCAT proteins, as is relevant to beta-catenin related disorders, such as cancer (e.g., for diagnostic means).

An MBCAT-interacting protein may be endogenous, i.e. one that naturally interacts genetically or biochemically with an MBCAT, such as a member of the MBCAT pathway that modulates MBCAT expression, localization, and/or activity. MBCAT-modulators include dominant negative forms of MBCAT-interacting proteins and of MBCAT proteins themselves. Yeast two-hybrid and variant screens offer preferred methods for identifying endogenous MBCAT-interacting proteins (Finley, R. L. et al. (1996) in DNA Cloning-Expression Systems: A Practical Approach, eds. Glover D. &

Hames B. D (Oxford University Press, Oxford, England), pp. 169-203; Fashema SF et al., *Gene* (2000) 250:1-14; Drees BL *Curr Opin Chem Biol* (1999) 3:64-70; Vidal M and Legrain P *Nucleic Acids Res* (1999) 27:919-29; and U.S. Pat. No. 5,928,868). Mass spectrometry is an alternative preferred method for the elucidation of protein complexes (reviewed in, e.g., Pandley A and Mann M, *Nature* (2000) 405:837-846; Yates JR 3rd, *Trends Genet* (2000) 16:5-8).

An MBCAT-interacting protein may be an exogenous protein, such as an MBCAT-specific antibody or a T-cell antigen receptor (see, e.g., Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory; Harlow and Lane (1999) *Using antibodies: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press). MBCAT antibodies are further discussed below.

In preferred embodiments, an MBCAT-interacting protein specifically binds an MBCAT protein. In alternative preferred embodiments, an MBCAT-modulating agent binds an MBCAT substrate, binding partner, or cofactor.

Antibodies

In another embodiment, the protein modulator is an MBCAT specific antibody agonist or antagonist. The antibodies have therapeutic and diagnostic utilities, and can be used in screening assays to identify MBCAT modulators. The antibodies can also be used in dissecting the portions of the MBCAT pathway responsible for various cellular responses and in the general processing and maturation of the MBCAT.

Antibodies that specifically bind MBCAT polypeptides can be generated using known methods. Preferably the antibody is specific to a mammalian ortholog of MBCAT polypeptide, and more preferably, to human MBCAT. Antibodies may be polyclonal, monoclonal (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab').sub.2 fragments, fragments produced by a FAb expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Epitopes of MBCAT which are particularly antigenic can be selected, for example, by routine screening of MBCAT polypeptides for antigenicity or by applying a theoretical method for selecting antigenic regions of a protein (Hopp and Wood (1981), *Proc. Natl. Acad. Sci. U.S.A.* 78:3824-28; Hopp and Wood, (1983) *Mol. Immunol.* 20:483-89;

Sutcliffe et al., (1983) Science 219:660-66) to the amino acid sequence of an MBCAT. Monoclonal antibodies with affinities of 10^8 M^{-1} preferably 10^9 M^{-1} to 10^{10} M^{-1} , or stronger can be made by standard procedures as described (Harlow and Lane, *supra*; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed) Academic Press, New York; and U.S. Pat. Nos. 4,381,292; 4,451,570; and 4,618,577). Antibodies may be generated against crude cell extracts of MBCAT or substantially purified fragments thereof. If MBCAT fragments are used, they preferably comprise at least 10, and more preferably, at least 20 contiguous amino acids of an MBCAT protein. In a particular embodiment, MBCAT-specific antigens and/or immunogens are coupled to carrier proteins that stimulate the immune response. For example, the subject polypeptides are covalently coupled to the keyhole limpet hemocyanin (KLH) carrier, and the conjugate is emulsified in Freund's complete adjuvant, which enhances the immune response. An appropriate immune system such as a laboratory rabbit or mouse is immunized according to conventional protocols.

The presence of MBCAT-specific antibodies is assayed by an appropriate assay such as a solid phase enzyme-linked immunosorbant assay (ELISA) using immobilized corresponding MBCAT polypeptides. Other assays, such as radioimmunoassays or fluorescent assays might also be used.

Chimeric antibodies specific to MBCAT polypeptides can be made that contain different portions from different animal species. For instance, a human immunoglobulin constant region may be linked to a variable region of a murine mAb, such that the antibody derives its biological activity from the human antibody, and its binding specificity from the murine fragment. Chimeric antibodies are produced by splicing together genes that encode the appropriate regions from each species (Morrison et al., Proc. Natl. Acad. Sci. (1984) 81:6851-6855; Neuberger et al., Nature (1984) 312:604-608; Takeda et al., Nature (1985) 31:452-454). Humanized antibodies, which are a form of chimeric antibodies, can be generated by grafting complementary-determining regions (CDRs) (Carlos, T. M., J. M. Harlan. 1994. Blood 84:2068-2101) of mouse antibodies into a background of human framework regions and constant regions by recombinant DNA technology (Riechmann LM, et al., 1988 Nature 323: 323-327). Humanized antibodies contain ~10% murine sequences and ~90% human sequences, and thus further

reduce or eliminate immunogenicity, while retaining the antibody specificities (Co MS, and Queen C. 1991 *Nature* 351: 501-501; Morrison SL. 1992 *Ann. Rev. Immun.* 10:239-265). Humanized antibodies and methods of their production are well-known in the art (U.S. Pat. Nos. 5,530,101, 5,585,089, 5,693,762, and 6,180,370).

MBCAT-specific single chain antibodies which are recombinant, single chain polypeptides formed by linking the heavy and light chain fragments of the Fv regions via an amino acid bridge, can be produced by methods known in the art (U.S. Pat. No. 4,946,778; Bird, *Science* (1988) 242:423-426; Huston et al., *Proc. Natl. Acad. Sci. USA* (1988) 85:5879-5883; and Ward et al., *Nature* (1989) 334:544-546).

Other suitable techniques for antibody production involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors (Huse et al., *Science* (1989) 246:1275-1281). As used herein, T-cell antigen receptors are included within the scope of antibody modulators (Harlow and Lane, 1988, *supra*).

The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, antibodies will be labeled by joining, either covalently or non-covalently, a substance that provides for a detectable signal, or that is toxic to cells that express the targeted protein (Menard S, et al., *Int J. Biol Markers* (1989) 4:131-134). A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, fluorescent emitting lanthanide metals, chemiluminescent moieties, bioluminescent moieties, magnetic particles, and the like (U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241). Also, recombinant immunoglobulins may be produced (U.S. Pat. No. 4,816,567). Antibodies to cytoplasmic polypeptides may be delivered and reach their targets by conjugation with membrane-penetrating toxin proteins (U.S. Pat. No. 6,086,900).

When used therapeutically in a patient, the antibodies of the subject invention are typically administered parenterally, when possible at the target site, or intravenously. The therapeutically effective dose and dosage regimen is determined by clinical studies. Typically, the amount of antibody administered is in the range of about 0.1 mg/kg –to

about 10 mg/kg of patient weight. For parenteral administration, the antibodies are formulated in a unit dosage injectable form (e.g., solution, suspension, emulsion) in association with a pharmaceutically acceptable vehicle. Such vehicles are inherently nontoxic and non-therapeutic. Examples are water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Nonaqueous vehicles such as fixed oils, ethyl oleate, or liposome carriers may also be used. The vehicle may contain minor amounts of additives, such as buffers and preservatives, which enhance isotonicity and chemical stability or otherwise enhance therapeutic potential. The antibodies' concentrations in such vehicles are typically in the range of about 1 mg/ml to about 10 mg/ml. Immunotherapeutic methods are further described in the literature (US Pat. No. 5,859,206; WO0073469).

Specific biotherapeutics

In a preferred embodiment, an MBCAT-interacting protein may have biotherapeutic applications. Biotherapeutic agents formulated in pharmaceutically acceptable carriers and dosages may be used to activate or inhibit signal transduction pathways. This modulation may be accomplished by binding a ligand, thus inhibiting the activity of the pathway; or by binding a receptor, either to inhibit activation of, or to activate, the receptor. Alternatively, the biotherapeutic may itself be a ligand capable of activating or inhibiting a receptor. Biotherapeutic agents and methods of producing them are described in detail in U.S. Pat. No. 6,146,628.

When the MBCAT is a ligand, it may be used as a biotherapeutic agent to activate or inhibit its natural receptor. Alternatively, antibodies against MBCAT, as described in the previous section, may be used as biotherapeutic agents.

When the MBCAT is a receptor, its ligand(s), antibodies to the ligand(s) or the MBCAT itself may be used as biotherapeutics to modulate the activity of MBCAT in the beta-catenin pathway.

Nucleic Acid Modulators

Other preferred MBCAT-modulating agents comprise nucleic acid molecules, such as antisense oligomers or double stranded RNA (dsRNA), which generally inhibit MBCAT

activity. Preferred nucleic acid modulators interfere with the function of the MBCAT nucleic acid such as DNA replication, transcription, translocation of the MBCAT RNA to the site of protein translation, translation of protein from the MBCAT RNA, splicing of the MBCAT RNA to yield one or more mRNA species, or catalytic activity which may be engaged in or facilitated by the MBCAT RNA.

In one embodiment, the antisense oligomer is an oligonucleotide that is sufficiently complementary to an MBCAT mRNA to bind to and prevent translation, preferably by binding to the 5' untranslated region. MBCAT-specific antisense oligonucleotides, preferably range from at least 6 to about 200 nucleotides. In some embodiments the oligonucleotide is preferably at least 10, 15, or 20 nucleotides in length. In other embodiments, the oligonucleotide is preferably less than 50, 40, or 30 nucleotides in length. The oligonucleotide can be DNA or RNA or a chimeric mixture or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may include other appending groups such as peptides, agents that facilitate transport across the cell membrane, hybridization-triggered cleavage agents, and intercalating agents.

In another embodiment, the antisense oligomer is a phosphothioate morpholino oligomer (PMO). PMOs are assembled from four different morpholino subunits, each of which contain one of four genetic bases (A, C, G, or T) linked to a six-membered morpholine ring. Polymers of these subunits are joined by non-ionic phosphodiarnidate intersubunit linkages. Details of how to make and use PMOs and other antisense oligomers are well known in the art (e.g. see WO99/18193; Probst JC, Antisense Oligodeoxynucleotide and Ribozyme Design, Methods. (2000) 22(3):271-281; Summerton J, and Weller D. 1997 Antisense Nucleic Acid Drug Dev. :7:187-95; US Pat. No. 5,235,033; and US Pat No. 5,378,841).

Alternative preferred MBCAT nucleic acid modulators are double-stranded RNA species mediating RNA interference (RNAi). RNAi is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene. Methods relating to the use of RNAi to silence genes in *C. elegans*, *Drosophila*, plants, and humans are

known in the art (Fire A, et al., 1998 Nature 391:806-811; Fire, A. Trends Genet. 15, 358-363 (1999); Sharp, P. A. RNA interference 2001. Genes Dev. 15, 485-490 (2001); Hammond, S. M., et al., Nature Rev. Genet. 2, 110-1119 (2001); Tuschl, T. Chem. Biochem. 2, 239-245 (2001); Hamilton, A. et al., Science 286, 950-952 (1999); Hammond, S. M., et al., Nature 404, 293-296 (2000); Zamore, P. D., et al., Cell 101, 25-33 (2000); Bernstein, E., et al., Nature 409, 363-366 (2001); Elbashir, S. M., et al., Genes Dev. 15, 188-200 (2001); WO0129058; WO9932619; Elbashir SM, et al., 2001 Nature 411:494-498).

Nucleic acid modulators are commonly used as research reagents, diagnostics, and therapeutics. For example, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used to elucidate the function of particular genes (see, for example, U.S. Pat. No. 6,165,790). Nucleic acid modulators are also used, for example, to distinguish between functions of various members of a biological pathway. For example, antisense oligomers have been employed as therapeutic moieties in the treatment of disease states in animals and man and have been demonstrated in numerous clinical trials to be safe and effective (Milligan JF, *et al*, Current Concepts in Antisense Drug Design, J Med Chem. (1993) 36:1923-1937; Tonkinson JL *et al*, Antisense Oligodeoxynucleotides as Clinical Therapeutic Agents, Cancer Invest. (1996) 14:54-65). Accordingly, in one aspect of the invention, an MBCAT-specific nucleic acid modulator is used in an assay to further elucidate the role of the MBCAT in the beta-catenin pathway, and/or its relationship to other members of the pathway. In another aspect of the invention, an MBCAT-specific antisense oligomer is used as a therapeutic agent for treatment of beta-catenin-related disease states.

Assay Systems

The invention provides assay systems and screening methods for identifying specific modulators of MBCAT activity. As used herein, an "assay system" encompasses all the components required for performing and analyzing results of an assay that detects and/or measures a particular event. In general, primary assays are used to identify or confirm a modulator's specific biochemical or molecular effect with respect to the MBCAT nucleic acid or protein. In general, secondary assays further assess the activity of a MBCAT

modulating agent identified by a primary assay and may confirm that the modulating agent affects MBCAT in a manner relevant to the beta-catenin pathway. In some cases, MBCAT modulators will be directly tested in a secondary assay.

In a preferred embodiment, the screening method comprises contacting a suitable assay system comprising an MBCAT polypeptide or nucleic acid with a candidate agent under conditions whereby, but for the presence of the agent, the system provides a reference activity (e.g. binding activity), which is based on the particular molecular event the screening method detects. A statistically significant difference between the agent-biased activity and the reference activity indicates that the candidate agent modulates MBCAT activity, and hence the beta-catenin pathway. The MBCAT polypeptide or nucleic acid used in the assay may comprise any of the nucleic acids or polypeptides described above.

Primary Assays

The type of modulator tested generally determines the type of primary assay.

Primary assays for small molecule modulators

For small molecule modulators, screening assays are used to identify candidate modulators. Screening assays may be cell-based or may use a cell-free system that recreates or retains the relevant biochemical reaction of the target protein (reviewed in Sittampalam GS *et al.*, Curr Opin Chem Biol (1997) 1:384-91 and accompanying references). As used herein the term "cell-based" refers to assays using live cells, dead cells, or a particular cellular fraction, such as a membrane, endoplasmic reticulum, or mitochondrial fraction. The term "cell free" encompasses assays using substantially purified protein (either endogenous or recombinantly produced), partially purified or crude cellular extracts. Screening assays may detect a variety of molecular events, including protein-DNA interactions, protein-protein interactions (*e.g.*, receptor-ligand binding), transcriptional activity (*e.g.*, using a reporter gene), enzymatic activity (*e.g.*, via a property of the substrate), activity of second messengers, immunogenicity and changes in cellular morphology or other cellular characteristics. Appropriate screening assays may use a wide range of detection methods including fluorescent, radioactive,

colorimetric, spectrophotometric, and amperometric methods, to provide a read-out for the particular molecular event detected.

Cell-based screening assays usually require systems for recombinant expression of MBCAT and any auxiliary proteins demanded by the particular assay. Appropriate methods for generating recombinant proteins produce sufficient quantities of proteins that retain their relevant biological activities and are of sufficient purity to optimize activity and assure assay reproducibility. Yeast two-hybrid and variant screens, and mass spectrometry provide preferred methods for determining protein-protein interactions and elucidation of protein complexes. In certain applications, when MBCAT-interacting proteins are used in screens to identify small molecule modulators, the binding specificity of the interacting protein to the MBCAT protein may be assayed by various known methods such as substrate processing (e.g. ability of the candidate MBCAT-specific binding agents to function as negative effectors in MBCAT-expressing cells), binding equilibrium constants (usually at least about 10^7 M^{-1} , preferably at least about 10^8 M^{-1} , more preferably at least about 10^9 M^{-1}), and immunogenicity (e.g. ability to elicit MBCAT specific antibody in a heterologous host such as a mouse, rat, goat or rabbit). For enzymes and receptors, binding may be assayed by, respectively, substrate and ligand processing.

The screening assay may measure a candidate agent's ability to specifically bind to or modulate activity of a MBCAT polypeptide, a fusion protein thereof, or to cells or membranes bearing the polypeptide or fusion protein. The MBCAT polypeptide can be full length or a fragment thereof that retains functional MBCAT activity. The MBCAT polypeptide may be fused to another polypeptide, such as a peptide tag for detection or anchoring, or to another tag. The MBCAT polypeptide is preferably human MBCAT, or is an ortholog or derivative thereof as described above. In a preferred embodiment, the screening assay detects candidate agent-based modulation of MBCAT interaction with a binding target, such as an endogenous or exogenous protein or other substrate that has MBCAT-specific binding activity, and can be used to assess normal MBCAT gene function.

Suitable assay formats that may be adapted to screen for MBCAT modulators are known in the art. Preferred screening assays are high throughput or ultra high throughput

and thus provide automated, cost-effective means of screening compound libraries for lead compounds (Fernandes PB, *Curr Opin Chem Biol* (1998) 2:597-603; Sundberg SA, *Curr Opin Biotechnol* 2000, 11:47-53). In one preferred embodiment, screening assays uses fluorescence technologies, including fluorescence polarization, time-resolved fluorescence, and fluorescence resonance energy transfer. These systems offer means to monitor protein-protein or DNA-protein interactions in which the intensity of the signal emitted from dye-labeled molecules depends upon their interactions with partner molecules (*e.g.*, Selvin PR, *Nat Struct Biol* (2000) 7:730-4; Fernandes PB, *supra*; Hertzberg RP and Pope AJ, *Curr Opin Chem Biol* (2000) 4:445-451).

A variety of suitable assay systems may be used to identify candidate MBCAT and beta-catenin pathway modulators (*e.g.* U.S. Pat. No. 6,165,992 (kinase assays); U.S. Pat. Nos. 5,550,019 and 6,133,437 (apoptosis assays); and U.S. Pat. Nos. 5,976,782, 6,225,118 and 6,444,434 (angiogenesis assays), among others). Specific preferred assays are described in more detail below.

Protein kinases, key signal transduction proteins that may be either membrane-associated or intracellular, catalyze the transfer of gamma phosphate from adenosine triphosphate (ATP) to a serine, threonine or tyrosine residue in a protein substrate. Radioassays, which monitor the transfer from [γ - ^{32}P or ^{33}P]ATP, are frequently used to assay kinase activity. For instance, a scintillation assay for p56 (lck) kinase activity monitors the transfer of the gamma phosphate from [γ - ^{33}P] ATP to a biotinylated peptide substrate. The substrate is captured on a streptavidin coated bead that transmits the signal (Beveridge M *et al.*, *J Biomol Screen* (2000) 5:205-212). This assay uses the scintillation proximity assay (SPA), in which only radio-ligand bound to receptors tethered to the surface of an SPA bead are detected by the scintillant immobilized within it, allowing binding to be measured without separation of bound from free ligand. Other assays for protein kinase activity may use antibodies that specifically recognize phosphorylated substrates. For instance, the kinase receptor activation (KIRA) assay measures receptor tyrosine kinase activity by ligand stimulating the intact receptor in cultured cells, then capturing solubilized receptor with specific antibodies and quantifying phosphorylation via phosphotyrosine ELISA (Sadick MD, *Dev Biol Stand* (1999) 97:121-133). Another example of antibody based assays for protein kinase

activity is TRF (time-resolved fluorometry). This method utilizes europium chelate-labeled anti-phosphotyrosine antibodies to detect phosphate transfer to a polymeric substrate coated onto microtiter plate wells. The amount of phosphorylation is then detected using time-resolved, dissociation-enhanced fluorescence (Braunwalder AF, et al., Anal Biochem 1996 Jul 1;238(2):159-64).

Transporter proteins carry a range of substrates, including nutrients, ions, amino acids, and drugs, across cell membranes. Assays for modulators of transporters may use labeled substrates. For instance, exemplary high throughput screens to identify compounds that interact with different peptide and anion transporters both use fluorescently labeled substrates; the assay for peptide transport additionally uses multiscreen filtration plates (Blevitt JM et al., J Biomol Screen 1999, 4:87-91; Cihlar T and Ho ES, Anal Biochem 2000, 283:49-55).

Assays for ATPase activity, such as described in Blackburn et al (Blackburn CL, et al., (1999) J Org Chem 64:5565-5570), are known in the art. The ATPase assay is performed using the EnzCheck ATPase kit (Molecular Probes). The assays are performed using an Ultraspec spectrophotometer (Pharmacia), and the progress of the reaction are monitored by absorbance increase at 360 nm. Microtubules (1.7 mM final), kinesin (0.11 mM final), inhibitor (or DMSO blank at 5% final), and the EnzCheck components (purine nucleotide phosphorylase and MESG substrate) are premixed in the cuvette in a reaction buffer (40 mM PIPES pH 6.8, 5 mM paclitaxel, 1 mM EGTA, 5 mM MgCl₂). The reaction is initiated by addition of MgATP (1 mM final).

Apoptosis assays. Assays for apoptosis may be performed by terminal deoxynucleotidyl transferase-mediated digoxigenin-11-dUTP nick end labeling (TUNEL) assay. The TUNEL assay is used to measure nuclear DNA fragmentation characteristic of apoptosis (Lazebnik *et al.*, 1994, Nature 371, 346), by following the incorporation of fluorescein-dUTP (Yonehara *et al.*, 1989, J. Exp. Med. 169, 1747). Apoptosis may further be assayed by acridine orange staining of tissue culture cells (Lucas, R., et al., 1998, Blood 15:4730-41). Other cell-based apoptosis assays include the caspase-3/7 assay and the cell death nucleosome ELISA assay. The caspase 3/7 assay is based on the activation of the caspase cleavage activity as part of a cascade of events that occur during

programmed cell death in many apoptotic pathways. In the caspase 3/7 assay (commercially available Apo-ONE™ Homogeneous Caspase-3/7 assay from Promega, cat# 67790), lysis buffer and caspase substrate are mixed and added to cells. The caspase substrate becomes fluorescent when cleaved by active caspase 3/7. The nucleosome ELISA assay is a general cell death assay known to those skilled in the art, and available commercially (Roche, Cat# 1774425). This assay is a quantitative sandwich-enzyme-immunoassay which uses monoclonal antibodies directed against DNA and histones respectively, thus specifically determining amount of mono- and oligonucleosomes in the cytoplasmic fraction of cell lysates. Mono and oligonucleosomes are enriched in the cytoplasm during apoptosis due to the fact that DNA fragmentation occurs several hours before the plasma membrane breaks down, allowing for accumulation in the cytoplasm. Nucleosomes are not present in the cytoplasmic fraction of cells that are not undergoing apoptosis. An apoptosis assay system may comprise a cell that expresses an MBCAT, and that optionally has defective beta-catenin function (e.g. beta-catenin is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the apoptosis assay system and changes in induction of apoptosis relative to controls where no test agent is added, identify candidate beta-catenin modulating agents. In some embodiments of the invention, an apoptosis assay may be used as a secondary assay to test a candidate beta-catenin modulating agents that is initially identified using a cell-free assay system. An apoptosis assay may also be used to test whether MBCAT function plays a direct role in apoptosis. For example, an apoptosis assay may be performed on cells that over- or under-express MBCAT relative to wild type cells. Differences in apoptotic response compared to wild type cells suggests that the MBCAT plays a direct role in the apoptotic response. Apoptosis assays are described further in US Pat. No. 6,133,437.

Cell proliferation and cell cycle assays. Cell proliferation may be assayed via bromodeoxyuridine (BRDU) incorporation. This assay identifies a cell population undergoing DNA synthesis by incorporation of BRDU into newly-synthesized DNA. Newly-synthesized DNA may then be detected using an anti-BRDU antibody (Hoshino *et*

al., 1986, *Int. J. Cancer* 38, 369; Campana *et al.*, 1988, *J. Immunol. Meth.* 107, 79), or by other means.

Cell proliferation is also assayed via phospho-histone H3 staining, which identifies a cell population undergoing mitosis by phosphorylation of histone H3. Phosphorylation of histone H3 at serine 10 is detected using an antibody specific to the phosphorylated form of the serine 10 residue of histone H3. (Chadlee, D.N. 1995, *J. Biol. Chem* 270:20098-105). Cell Proliferation may also be examined using [^3H]-thymidine incorporation (Chen, J., 1996, *Oncogene* 13:1395-403; Jeoung, J., 1995, *J. Biol. Chem.* 270:18367-73). This assay allows for quantitative characterization of S-phase DNA syntheses. In this assay, cells synthesizing DNA will incorporate [^3H]-thymidine into newly synthesized DNA. Incorporation can then be measured by standard techniques such as by counting of radioisotope in a scintillation counter (e.g., Beckman LS 3800 Liquid Scintillation Counter). Another proliferation assay uses the dye Alamar Blue (available from Biosource International), which fluoresces when reduced in living cells and provides an indirect measurement of cell number (Voytik-Harbin SL *et al.*, 1998, *In Vitro Cell Dev Biol Anim* 34:239-46). Yet another proliferation assay, the MTS assay, is based on in vitro cytotoxicity assessment of industrial chemicals, and uses the soluble tetrazolium salt, MTS. MTS assays are commercially available, for example, the Promega CellTiter 96[®] Aqueous Non-Radioactive Cell Proliferation Assay (Cat.# G5421).

Cell proliferation may also be assayed by colony formation in soft agar (Sambrook *et al.*, *Molecular Cloning*, Cold Spring Harbor (1989)). For example, cells transformed with MBCAT are seeded in soft agar plates, and colonies are measured and counted after two weeks incubation.

Cell proliferation may also be assayed by measuring ATP levels as indicator of metabolically active cells. Such assays are commercially available, for example Cell Titer-Glo[™], which is a luminescent homogeneous assay available from Promega.

Involvement of a gene in the cell cycle may be assayed by flow cytometry (Gray JW *et al.* (1986) *Int J Radiat Biol Relat Stud Phys Chem Med* 49:237-55). Cells transfected with an MBCAT may be stained with propidium iodide and evaluated in a flow cytometer (available from Becton Dickinson), which indicates accumulation of cells in different stages of the cell cycle.

Accordingly, a cell proliferation or cell cycle assay system may comprise a cell that expresses an MBCAT, and that optionally has defective beta-catenin function (e.g. beta-catenin is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the assay system and changes in cell proliferation or cell cycle relative to controls where no test agent is added, identify candidate beta-catenin modulating agents. In some embodiments of the invention, the cell proliferation or cell cycle assay may be used as a secondary assay to test a candidate beta-catenin modulating agents that is initially identified using another assay system such as a cell-free assay system. A cell proliferation assay may also be used to test whether MBCAT function plays a direct role in cell proliferation or cell cycle. For example, a cell proliferation or cell cycle assay may be performed on cells that over- or under-express MBCAT relative to wild type cells. Differences in proliferation or cell cycle compared to wild type cells suggests that the MBCAT plays a direct role in cell proliferation or cell cycle.

Angiogenesis. Angiogenesis may be assayed using various human endothelial cell systems, such as umbilical vein, coronary artery, or dermal cells. Suitable assays include Alamar Blue based assays (available from Biosource International) to measure proliferation; migration assays using fluorescent molecules, such as the use of Becton Dickinson Falcon HTS FluoroBlock cell culture inserts to measure migration of cells through membranes in presence or absence of angiogenesis enhancer or suppressors; and tubule formation assays based on the formation of tubular structures by endothelial cells on Matrigel® (Becton Dickinson). Accordingly, an angiogenesis assay system may comprise a cell that expresses an MBCAT, and that optionally has defective beta-catenin function (e.g. beta-catenin is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the angiogenesis assay system and changes in angiogenesis relative to controls where no test agent is added, identify candidate beta-catenin modulating agents. In some embodiments of the invention, the angiogenesis assay may be used as a secondary assay to test a candidate beta-catenin modulating agents that is initially identified using another assay system. An angiogenesis assay may also be used to test whether MBCAT function plays a direct role in cell proliferation. For example, an angiogenesis assay may be performed on cells that over- or under-express

MBCAT relative to wild type cells. Differences in angiogenesis compared to wild type cells suggests that the MBCAT plays a direct role in angiogenesis. U.S. Pat. Nos. 5,976,782, 6,225,118 and 6,444,434, among others, describe various angiogenesis assays.

Hypoxic induction. The alpha subunit of the transcription factor, hypoxia inducible factor-1 (HIF-1), is upregulated in tumor cells following exposure to hypoxia in vitro. Under hypoxic conditions, HIF-1 stimulates the expression of genes known to be important in tumour cell survival, such as those encoding glycolytic enzymes and VEGF. Induction of such genes by hypoxic conditions may be assayed by growing cells transfected with MBCAT in hypoxic conditions (such as with 0.1% O₂, 5% CO₂, and balance N₂, generated in a Napco 7001 incubator (Precision Scientific)) and normoxic conditions, followed by assessment of gene activity or expression by Taqman®. For example, a hypoxic induction assay system may comprise a cell that expresses an MBCAT, and that optionally has defective beta-catenin function (e.g. beta-catenin is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the hypoxic induction assay system and changes in hypoxic response relative to controls where no test agent is added, identify candidate beta-catenin modulating agents. In some embodiments of the invention, the hypoxic induction assay may be used as a secondary assay to test a candidate beta-catenin modulating agents that is initially identified using another assay system. A hypoxic induction assay may also be used to test whether MBCAT function plays a direct role in the hypoxic response. For example, a hypoxic induction assay may be performed on cells that over- or under-express MBCAT relative to wild type cells. Differences in hypoxic response compared to wild type cells suggests that the MBCAT plays a direct role in hypoxic induction.

Cell adhesion. Cell adhesion assays measure adhesion of cells to purified adhesion proteins, or adhesion of cells to each other, in presence or absence of candidate modulating agents. Cell-protein adhesion assays measure the ability of agents to modulate the adhesion of cells to purified proteins. For example, recombinant proteins are produced, diluted to 2.5g/mL in PBS, and used to coat the wells of a microtiter plate. The wells used for negative control are not coated. Coated wells are then washed,

blocked with 1% BSA, and washed again. Compounds are diluted to 2× final test concentration and added to the blocked, coated wells. Cells are then added to the wells, and the unbound cells are washed off. Retained cells are labeled directly on the plate by adding a membrane-permeable fluorescent dye, such as calcein-AM, and the signal is quantified in a fluorescent microplate reader.

Cell-cell adhesion assays measure the ability of agents to modulate binding of cell adhesion proteins with their native ligands. These assays use cells that naturally or recombinantly express the adhesion protein of choice. In an exemplary assay, cells expressing the cell adhesion protein are plated in wells of a multiwell plate. Cells expressing the ligand are labeled with a membrane-permeable fluorescent dye, such as BCECF, and allowed to adhere to the monolayers in the presence of candidate agents. Unbound cells are washed off, and bound cells are detected using a fluorescence plate reader.

High-throughput cell adhesion assays have also been described. In one such assay, small molecule ligands and peptides are bound to the surface of microscope slides using a microarray spotter, intact cells are then contacted with the slides, and unbound cells are washed off. In this assay, not only the binding specificity of the peptides and modulators against cell lines are determined, but also the functional cell signaling of attached cells using immunofluorescence techniques in situ on the microchip is measured (Falsey JR et al., *Bioconjug Chem.* 2001 May-Jun;12(3):346-53).

Tubulogenesis. Tubulogenesis assays monitor the ability of cultured cells, generally endothelial cells, to form tubular structures on a matrix substrate, which generally simulates the environment of the extracellular matrix. Exemplary substrates include Matrigel™ (Becton Dickinson), an extract of basement membrane proteins containing laminin, collagen IV, and heparin sulfate proteoglycan, which is liquid at 4°C and forms a solid gel at 37°C. Other suitable matrices comprise extracellular components such as collagen, fibronectin, and/or fibrin. Cells are stimulated with a pro-angiogenic stimulant, and their ability to form tubules is detected by imaging. Tubules can generally be detected after an overnight incubation with stimuli, but longer or shorter time frames may also be used. Tube formation assays are well known in the art (e.g.,

Jones MK et al., 1999, *Nature Medicine* 5:1418-1423). These assays have traditionally involved stimulation with serum or with the growth factors FGF or VEGF. Serum represents an undefined source of growth factors. In a preferred embodiment, the assay is performed with cells cultured in serum free medium, in order to control which process or pathway a candidate agent modulates. Moreover, we have found that different target genes respond differently to stimulation with different pro-angiogenic agents, including inflammatory angiogenic factors such as TNF- α . Thus, in a further preferred embodiment, a tubulogenesis assay system comprises testing an MBCAT's response to a variety of factors, such as FGF, VEGF, phorbol myristate acetate (PMA), TNF- α , ephrin, etc.

Cell Migration. An invasion/migration assay (also called a migration assay) tests the ability of cells to overcome a physical barrier and to migrate towards pro-angiogenic signals. Migration assays are known in the art (e.g., Paik JH et al., 2001, *J Biol Chem* 276:11830-11837). In a typical experimental set-up, cultured endothelial cells are seeded onto a matrix-coated porous lamina, with pore sizes generally smaller than typical cell size. The matrix generally simulates the environment of the extracellular matrix, as described above. The lamina is typically a membrane, such as the transwell polycarbonate membrane (Corning Costar Corporation, Cambridge, MA), and is generally part of an upper chamber that is in fluid contact with a lower chamber containing pro-angiogenic stimuli. Migration is generally assayed after an overnight incubation with stimuli, but longer or shorter time frames may also be used. Migration is assessed as the number of cells that crossed the lamina, and may be detected by staining cells with hemotoxylin solution (VWR Scientific, South San Francisco, CA), or by any other method for determining cell number. In another exemplary set up, cells are fluorescently labeled and migration is detected using fluorescent readings, for instance using the Falcon HTS FluoroBlok (Becton Dickinson). While some migration is observed in the absence of stimulus, migration is greatly increased in response to pro-angiogenic factors. As described above, a preferred assay system for migration/invasion assays comprises testing an MBCAT's response to a variety of pro-angiogenic factors,

including tumor angiogenic and inflammatory angiogenic agents, and culturing the cells in serum free medium.

Sprouting assay. A sprouting assay is a three-dimensional *in vitro* angiogenesis assay that uses a cell-number defined spheroid aggregation of endothelial cells ("spheroid"), embedded in a collagen gel-based matrix. The spheroid can serve as a starting point for the sprouting of capillary-like structures by invasion into the extracellular matrix (termed "cell sprouting") and the subsequent formation of complex anastomosing networks (Korff and Augustin, 1999, J Cell Sci 112:3249-58). In an exemplary experimental set-up, spheroids are prepared by pipetting 400 human umbilical vein endothelial cells into individual wells of a nonadhesive 96-well plates to allow overnight spheroidal aggregation (Korff and Augustin: J Cell Biol 143: 1341-52, 1998). Spheroids are harvested and seeded in 900 μ l of methocel-collagen solution and pipetted into individual wells of a 24 well plate to allow collagen gel polymerization. Test agents are added after 30 min by pipetting 100 μ l of 10-fold concentrated working dilution of the test substances on top of the gel. Plates are incubated at 37°C for 24h. Dishes are fixed at the end of the experimental incubation period by addition of paraformaldehyde. Sprouting intensity of endothelial cells can be quantitated by an automated image analysis system to determine the cumulative sprout length per spheroid.

Primary assays for antibody modulators

For antibody modulators, appropriate primary assays test is a binding assay that tests the antibody's affinity to and specificity for the MBCAT protein. Methods for testing antibody affinity and specificity are well known in the art (Harlow and Lane, 1988, 1999, *supra*). The enzyme-linked immunosorbant assay (ELISA) is a preferred method for detecting MBCAT-specific antibodies; others include FACS assays, radioimmunoassays, and fluorescent assays.

In some cases, screening assays described for small molecule modulators may also be used to test antibody modulators.

Primary assays for nucleic acid modulators

For nucleic acid modulators, primary assays may test the ability of the nucleic acid modulator to inhibit or enhance MBCAT gene expression, preferably mRNA expression. In general, expression analysis comprises comparing MBCAT expression in like populations of cells (*e.g.*, two pools of cells that endogenously or recombinantly express MBCAT) in the presence and absence of the nucleic acid modulator. Methods for analyzing mRNA and protein expression are well known in the art. For instance, Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR (*e.g.*, using the TaqMan®, PE Applied Biosystems), or microarray analysis may be used to confirm that MBCAT mRNA expression is reduced in cells treated with the nucleic acid modulator (*e.g.*, Current Protocols in Molecular Biology (1994) Ausubel FM *et al.*, *eds.*, John Wiley & Sons, Inc., chapter 4; Freeman WM *et al.*, Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann Med 2001, 33:142-147; Blohm DH and Guiseppi-Elie, A Curr Opin Biotechnol 2001, 12:41-47). Protein expression may also be monitored. Proteins are most commonly detected with specific antibodies or antisera directed against either the MBCAT protein or specific peptides. A variety of means including Western blotting, ELISA, or in situ detection, are available (Harlow E and Lane D, 1988 and 1999, *supra*).

In some cases, screening assays described for small molecule modulators, particularly in assay systems that involve MBCAT mRNA expression, may also be used to test nucleic acid modulators.

Secondary Assays

Secondary assays may be used to further assess the activity of MBCAT-modulating agent identified by any of the above methods to confirm that the modulating agent affects MBCAT in a manner relevant to the beta-catenin pathway. As used herein, MBCAT-modulating agents encompass candidate clinical compounds or other agents derived from previously identified modulating agent. Secondary assays can also be used to test the activity of a modulating agent on a particular genetic or biochemical pathway or to test the specificity of the modulating agent's interaction with MBCAT.

Secondary assays generally compare like populations of cells or animals (*e.g.*, two pools of cells or animals that endogenously or recombinantly express MBCAT) in the presence and absence of the candidate modulator. In general, such assays test whether

treatment of cells or animals with a candidate MBCAT–modulating agent results in changes in the beta-catenin pathway in comparison to untreated (or mock- or placebo-treated) cells or animals. Certain assays use “sensitized genetic backgrounds”, which, as used herein, describe cells or animals engineered for altered expression of genes in the beta-catenin or interacting pathways.

Cell-based assays

Cell based assays may detect endogenous beta-catenin pathway activity or may rely on recombinant expression of beta-catenin pathway components. Any of the aforementioned assays may be used in this cell-based format. Candidate modulators are typically added to the cell media but may also be injected into cells or delivered by any other efficacious means.

Animal Assays

A variety of non-human animal models of normal or defective beta-catenin pathway may be used to test candidate MBCAT modulators. Models for defective beta-catenin pathway typically use genetically modified animals that have been engineered to mis-express (*e.g.*, over-express or lack expression in) genes involved in the beta-catenin pathway. Assays generally require systemic delivery of the candidate modulators, such as by oral administration, injection, etc.

In a preferred embodiment, beta-catenin pathway activity is assessed by monitoring neovascularization and angiogenesis. Animal models with defective and normal beta-catenin are used to test the candidate modulator’s affect on MBCAT in Matrigel® assays. Matrigel® is an extract of basement membrane proteins, and is composed primarily of laminin, collagen IV, and heparin sulfate proteoglycan. It is provided as a sterile liquid at 4° C, but rapidly forms a solid gel at 37° C. Liquid Matrigel® is mixed with various angiogenic agents, such as bFGF and VEGF, or with human tumor cells which over-express the MBCAT. The mixture is then injected subcutaneously(SC) into female athymic nude mice (Taconic, Germantown, NY) to support an intense vascular response. Mice with Matrigel® pellets may be dosed via oral (PO), intraperitoneal (IP), or intravenous (IV) routes with the candidate modulator. Mice are euthanized 5 - 12 days

post-injection, and the Matrigel® pellet is harvested for hemoglobin analysis (Sigma plasma hemoglobin kit). Hemoglobin content of the gel is found to correlate the degree of neovascularization in the gel.

In another preferred embodiment, the effect of the candidate modulator on MBCAT is assessed via tumorigenicity assays. Tumor xenograft assays are known in the art (see, e.g., Ogawa K et al., 2000, *Oncogene* 19:6043-6052). Xenografts are typically implanted SC into female athymic mice, 6-7 week old, as single cell suspensions either from a pre-existing tumor or from *in vitro* culture. The tumors which express the MBCAT endogenously are injected in the flank, 1×10^5 to 1×10^7 cells per mouse in a volume of 100 μ L using a 27gauge needle. Mice are then ear tagged and tumors are measured twice weekly. Candidate modulator treatment is initiated on the day the mean tumor weight reaches 100 mg. Candidate modulator is delivered IV, SC, IP, or PO by bolus administration. Depending upon the pharmacokinetics of each unique candidate modulator, dosing can be performed multiple times per day. The tumor weight is assessed by measuring perpendicular diameters with a caliper and calculated by multiplying the measurements of diameters in two dimensions. At the end of the experiment, the excised tumors may be utilized for biomarker identification or further analyses. For immunohistochemistry staining, xenograft tumors are fixed in 4% paraformaldehyde, 0.1M phosphate, pH 7.2, for 6 hours at 4°C, immersed in 30% sucrose in PBS, and rapidly frozen in isopentane cooled with liquid nitrogen.

In another preferred embodiment, tumorigenicity is monitored using a hollow fiber assay, which is described in U.S. Pat No. US 5,698,413. Briefly, the method comprises implanting into a laboratory animal a biocompatible, semi-permeable encapsulation device containing target cells, treating the laboratory animal with a candidate modulating agent, and evaluating the target cells for reaction to the candidate modulator. Implanted cells are generally human cells from a pre-existing tumor or a tumor cell line. After an appropriate period of time, generally around six days, the implanted samples are harvested for evaluation of the candidate modulator. Tumorigenicity and modulator efficacy may be evaluated by assaying the quantity of viable cells present in the macrocapsule, which can be determined by tests known in the art, for example, MTT dye conversion assay, neutral red dye uptake, trypan blue staining,

viable cell counts, the number of colonies formed in soft agar, the capacity of the cells to recover and replicate in vitro, etc.

In another preferred embodiment, a tumorigenicity assay use a transgenic animal, usually a mouse, carrying a dominant oncogene or tumor suppressor gene knockout under the control of tissue specific regulatory sequences; these assays are generally referred to as transgenic tumor assays. In a preferred application, tumor development in the transgenic model is well characterized or is controlled. In an exemplary model, the "RIP1-Tag2" transgene, comprising the SV40 large T-antigen oncogene under control of the insulin gene regulatory regions is expressed in pancreatic beta cells and results in islet cell carcinomas (Hanahan D, 1985, Nature 315:115-122; Parangi S et al, 1996, Proc Natl Acad Sci USA 93: 2002-2007; Bergers G et al, 1999, Science 284:808-812). An "angiogenic switch," occurs at approximately five weeks, as normally quiescent capillaries in a subset of hyperproliferative islets become angiogenic. The RIP1-TAG2 mice die by age 14 weeks. Candidate modulators may be administered at a variety of stages, including just prior to the angiogenic switch (e.g., for a model of tumor prevention), during the growth of small tumors (e.g., for a model of intervention), or during the growth of large and/or invasive tumors (e.g., for a model of regression). Tumorigenicity and modulator efficacy can be evaluating life-span extension and/or tumor characteristics, including number of tumors, tumor size, tumor morphology, vessel density, apoptotic index, etc.

Diagnostic and therapeutic uses

Specific MBCAT-modulating agents are useful in a variety of diagnostic and therapeutic applications where disease or disease prognosis is related to defects in the beta-catenin pathway, such as angiogenic, apoptotic, or cell proliferation disorders. Accordingly, the invention also provides methods for modulating the beta-catenin pathway in a cell, preferably a cell pre-determined to have defective or impaired beta-catenin function (e.g. due to overexpression, underexpression, or misexpression of beta-catenin, or due to gene mutations), comprising the step of administering an agent to the cell that specifically modulates MBCAT activity. Preferably, the modulating agent produces a detectable phenotypic change in the cell indicating that the beta-catenin

function is restored. The phrase "function is restored", and equivalents, as used herein, means that the desired phenotype is achieved, or is brought closer to normal compared to untreated cells. For example, with restored beta-catenin function, cell proliferation and/or progression through cell cycle may normalize, or be brought closer to normal relative to untreated cells. The invention also provides methods for treating disorders or disease associated with impaired beta-catenin function by administering a therapeutically effective amount of an MBCAT -modulating agent that modulates the beta-catenin pathway. The invention further provides methods for modulating MBCAT function in a cell, preferably a cell pre-determined to have defective or impaired MBCAT function, by administering an MBCAT -modulating agent. Additionally, the invention provides a method for treating disorders or disease associated with impaired MBCAT function by administering a therapeutically effective amount of an MBCAT -modulating agent.

The discovery that MBCAT is implicated in beta-catenin pathway provides for a variety of methods that can be employed for the diagnostic and prognostic evaluation of diseases and disorders involving defects in the beta-catenin pathway and for the identification of subjects having a predisposition to such diseases and disorders.

Various expression analysis methods can be used to diagnose whether MBCAT expression occurs in a particular sample, including Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR, and microarray analysis. (*e.g.*, Current Protocols in Molecular Biology (1994) Ausubel FM *et al.*, eds., John Wiley & Sons, Inc., chapter 4; Freeman WM *et al.*, Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann Med 2001, 33:142-147; Blohm and Guiseppi-Elie, Curr Opin Biotechnol 2001, 12:41-47). Tissues having a disease or disorder implicating defective beta-catenin signaling that express an MBCAT, are identified as amenable to treatment with an MBCAT modulating agent. In a preferred application, the beta-catenin defective tissue overexpresses an MBCAT relative to normal tissue. For example, a Northern blot analysis of mRNA from tumor and normal cell lines, or from tumor and matching normal tissue samples from the same patient, using full or partial MBCAT cDNA sequences as probes, can determine whether particular tumors express or overexpress MBCAT. Alternatively, the TaqMan® is used for quantitative RT-PCR analysis of MBCAT expression in cell lines, normal tissues and tumor samples (PE Applied Biosystems).

Various other diagnostic methods may be performed, for example, utilizing reagents such as the MBCAT oligonucleotides, and antibodies directed against an MBCAT, as described above for: (1) the detection of the presence of MBCAT gene mutations, or the detection of either over- or under-expression of MBCAT mRNA relative to the non-disorder state; (2) the detection of either an over- or an under-abundance of MBCAT gene product relative to the non-disorder state; and (3) the detection of perturbations or abnormalities in the signal transduction pathway mediated by MBCAT.

Thus, in a specific embodiment, the invention is drawn to a method for diagnosing a disease or disorder in a patient that is associated with alterations in MBCAT expression, the method comprising: a) obtaining a biological sample from the patient; b) contacting the sample with a probe for MBCAT expression; c) comparing results from step (b) with a control; and d) determining whether step (c) indicates a likelihood of the disease or disorder. Preferably, the disease is cancer. The probe may be either DNA or protein, including an antibody.

EXAMPLES

The following experimental section and examples are offered by way of illustration and not by way of limitation.

I. *Drosophila* beta-catenin screens

Two dominant loss of function screens were carried out in *Drosophila* to identify genes that interact with the Wg cell signaling molecule, beta –catenin (Riggleman et al. (1990) Cell 63:549-560; Peifer et al. (1991) Development 111:1029-1043). Late stage activation of the pathway in the developing *Drosophila* eye leads to apoptosis (Freeman and Bienz (2001) EMBO reports 2: 157-162), whereas early stage activation leads to an overgrowth phenotype. We discovered that ectopic expression of the activated protein in the wing results in changes of cell fate into ectopic bristles and wing veins.

Each transgene was carried in a separate fly stock:

Stocks and genotypes were as follows:

eye overgrowth transgene: isow; P{3.5 eyeless-Gal4}; P{arm(S56F)-pExp-UAS))/TM6b;

eye apoptosis transgene: *y w*; P{arm(S56F)-pExp-GMR}/CyO; and

wing transgene: P{arm(Δ N)-pExp-VgMQ}/FM7c

In the first dominant loss of function screen, females of each of these three transgenes were crossed to a collection of males containing genomic deficiencies. Resulting progeny containing the transgene and the deficiency were then scored for the effect of the deficiency on the eye apoptosis, eye overgrowth, and wing phenotypes, i.e., whether the deficiency enhanced, suppressed, or had no effect on their respective phenotypes. All data was recorded and all modifiers were retested with a repeat of the original cross. Modifying deficiencies of the phenotypes were then prioritized according to how they modified each of the three phenotypes.

Transposons contained within the prioritized deficiencies were then screened as described. Females of each of the three transgenes were crossed to a collection of 4 types of transposons (3 piggyBac-based and 1 P-element-based). The resulting progeny containing the transgene and the transposon were scored for the effect of the transposon on their respective phenotypes. All data was recorded and all modifiers were retested with a repeat of the original cross. Modifiers of the phenotypes were identified as either members of the Wg pathway, components of apoptotic related pathways, components of cell cycle related pathways, or cell adhesion related proteins.

In the second dominant loss of function screen, females of the eye overgrowth transgene were crossed to males from a collection of 3 types of piggyBac-based transposons. The resulting progeny containing the transgene and the transposon were scored for the effect of the transposon on the eye overgrowth phenotype. All data was recorded and all modifiers were retested with a repeat of the original cross. Modifiers of the phenotypes were identified as either members of the Wg pathway, components of cell cycle related pathways, or cell adhesion related proteins.

II. Analysis of Table 1

BLAST analysis (Altschul et al., *supra*) was employed to identify orthologs of *Drosophila* modifiers. The columns "MBCAT symbol", and "MBCAT name aliases " provide a symbol and the known name abbreviations for the Targets, where available, from Genbank. "MBCAT RefSeq_NA or GI_NA", "MBCAT GI_AA", "MBCAT

NAME”, and “MBCAT Description” provide the reference DNA sequences for the MBCATs as available from National Center for Biology Information (NCBI), MBCAT protein Genbank identifier number (GI#), MBCAT name, and MBCAT description, all available from Genbank, respectively. The length of each amino acid is in the “MBCAT Protein Length” column.

Names and Protein sequences of *Drosophila* modifiers of beta-catenin from screen (Example I), are represented in the “Modifier Name” and “Modifier GI_AA” column by GI#, respectively.

Table 1

MBCAT symbol	MBCAT name aliases	MBCAT RefSeq_NA or GI_NA	MBCAT GI_AA	MBCAT name	MBCAT description	MBCAT protein length	Modifier name	Modifier gi_aa
BAIAP3	BAIAP3 BAP3 KIAA0734 BAI-associated protein 3 BAI1-associated protein 3	NM_003933	15451914	BAI1-associated protein 3	protein C-terminus binding	1187	unc-13-4A	28574960
UNC13D	LOC201294 Munc13-4 protein Munc13-4 unc-13 homolog D (C. elegans) UNC13D	XM_113950	27500274	unc-13 homolog D (C. elegans)	na	1090	unc-13-4A	28574960
ABCC4	ABCC4 MRP4 MOATB MOAT-B EST170205 canalicular multispecific organic anion transporter (ABC superfamily) ATP-binding cassette, sub-family C (CFTR/MRP), member 4	NM_005845	5031915	ATP-binding cassette, sub-family C (CFTR/MRP), member 4	anion transporter; nucleotide transporter; ATP-binding cassette (ABC) transporter; multidrug efflux pump; methotrexate transporter	1325	CG9270	22946960

C14orf35	C14orf35 c14_5368 chromosome 14 open reading frame 35	XM_058661	18596951	chromosome 14 open reading frame 35	na	362	CG11412	24639058
CGI-130	CGI-130 NS5ATP2 CGI-130 protein	NM_016063	33328294	CGI-130 protein	na	204	CG11050	20129283
GSTZ1	GSTZ1 MAI MAAI GSTZ1-1 MGC2029 maleylacetone isomerase maleylacetoacetate isomerase glutathione S-aryltransferase glutathione S-alkyltransferase glutathione S-aralkyltransferase glutathione s-transferase Zeta 1 S-(hydroxyalkyl)glutathione lyase glutathione transferase zeta 1 (maleylacetoacetate isomerase)	NM_001513 NM_145870 NM_145871	22202624	glutathione transferase zeta 1 (maleylacetoacetate isomerase)	glutathione transferase; glutathione transferase; maleylacetoacetate isomerase; maleylacetoacetate isomerase; maleylacetoacetate isomerase; maleylacetoacetate isomerase; glutathione peroxidase	216	CG9363	24645375
STK18	SAK Snk Akin Kinase serine/threonine kinase 18 PLK2 STK18 Sak PLK4	NM_014264	21361433	serine/threonine kinase 18	protein serine/threonine kinase	970	SAK	21358073
TTBK1	KIAA1855 protein KIAA1855 TTBK1 tau tubulin kinase 1	XM_166453	20555151	tau tubulin kinase 1	protein kinase; ATP binding; casein kinase I	1270	CG11533	24638629
TTBK2	similar to tau-tubulin kinase LOC146057 tau-tubulin kinase KIAA0847 TTBK2 TTBK tau tubulin kinase 2	NM_173500	28466991	tau tubulin kinase 2	na	1649	CG11533	24638629

NT5E	NT5E eN NT5 NTE eNT CD73 E5NT ecto-5'- nucleotidase 5'- nucleotidase (CD73) Purine 5-Prime- Nucleotidase 5'-nucleotidase, ecto (CD73)	NM_002526	4505467	5'-nucleotidase, ecto (CD73)	cell adhesion molecule; 5'- nucleotidase; 5'- nucleotidase; 5'- nucleotidase	574	CG301 04	246544 24
RAD51 C	RAD51C RAD51L2 yeast RAD51 homolog 3 DNA repair protein RAD51 homolog 3 RAD51 homolog C (S. cerevisiae)	NM_002876 N M_058216 N M_058217	17402896	RAD51 homolog C (S. cerevisiae)	DNA binding; single-stranded DNA binding; ATPase; ATPase; protein binding; protein binding; double-stranded DNA binding; damaged DNA binding	376	spn-D	231800 23
XRCC 2	XRCC2 DNA repair protein XRCC2 X-ray repair, complementing defective, repair in Chinese hamster X-ray repair complementing defective repair in Chinese hamster cells 2	NM_005431	4885657	X-ray repair complementing defective repair in Chinese hamster cells 2	ATPase	280	spn-D	231800 23
SHANK1	SHANK1 SSTRIP SPANK-1 synamon somatostatin receptor- interacting protein SH3 and multiple ankyrin repeat domains 1	NM_016148	11968152	SH3 and multiple ankyrin repeat domains 1	cytoskeletal anchor protein; cytoskeletal protein binding	2161	Prosap	246535 38

SHANK2	SHANK2 SHANK CORTBP1 ProSAP1 SPANK-3 KIAA1022 cortactin binding protein 1 GKAP/SAPAP interacting protein cortactin SH3 domain-binding protein SH3 and multiple ankyrin repeat domains 2	NM_012309 M_133266	19743794	SH3 and multiple ankyrin repeat domains 2	na	1254	Prosap	24653538
SHANK3	SHANK3 SHANK3 PSAP2 PROSAP2 SPANK-2 KIAA1650 shank postsynaptic density protein proline rich synapse associated protein 2 (rat) SH3 and multiple ankyrin repeat domains 3	XM_037493	37564792	SH3 and multiple ankyrin repeat domains 3	cytoskeletal anchor protein	1784	Prosap	24653538

III. High-Throughput In Vitro Fluorescence Polarization Assay

Fluorescently-labeled MBCAT peptide/substrate are added to each well of a 96-well microtiter plate, along with a test agent in a test buffer (10 mM HEPES, 10 mM NaCl, 6 mM magnesium chloride, pH 7.6). Changes in fluorescence polarization, determined by using a Fluorolite FPM-2 Fluorescence Polarization Microtiter System (Dynatech Laboratories, Inc), relative to control values indicates the test compound is a candidate modifier of MBCAT activity.

IV. High-Throughput In Vitro Binding Assay.

³³P-labeled MBCAT peptide is added in an assay buffer (100 mM KCl, 20 mM HEPES pH 7.6, 1 mM MgCl₂, 1% glycerol, 0.5% NP-40, 50 mM beta-mercaptoethanol, 1

mg/ml BSA, cocktail of protease inhibitors) along with a test agent to the wells of a Neutralite-avidin coated assay plate and incubated at 25°C for 1 hour. Biotinylated substrate is then added to each well and incubated for 1 hour. Reactions are stopped by washing with PBS, and counted in a scintillation counter. Test agents that cause a difference in activity relative to control without test agent are identified as candidate beta-catenin modulating agents.

V. Immunoprecipitations and Immunoblotting

For coprecipitation of transfected proteins, 3×10^6 appropriate recombinant cells containing the MBCAT proteins are plated on 10-cm dishes and transfected on the following day with expression constructs. The total amount of DNA is kept constant in each transfection by adding empty vector. After 24 h, cells are collected, washed once with phosphate-buffered saline and lysed for 20 min on ice in 1 ml of lysis buffer containing 50 mM Hepes, pH 7.9, 250 mM NaCl, 20 mM -glycerophosphate, 1 mM sodium orthovanadate, 5 mM p-nitrophenyl phosphate, 2 mM dithiothreitol, protease inhibitors (complete, Roche Molecular Biochemicals), and 1% Nonidet P-40. Cellular debris is removed by centrifugation twice at $15,000 \times g$ for 15 min. The cell lysate is incubated with 25 μ l of M2 beads (Sigma) for 2 h at 4 °C with gentle rocking.

After extensive washing with lysis buffer, proteins bound to the beads are solubilized by boiling in SDS sample buffer, fractionated by SDS-polyacrylamide gel electrophoresis, transferred to polyvinylidene difluoride membrane and blotted with the indicated antibodies. The reactive bands are visualized with horseradish peroxidase coupled to the appropriate secondary antibodies and the enhanced chemiluminescence (ECL) Western blotting detection system (Amersham Pharmacia Biotech).

VI. Kinase assay

A purified or partially purified MBCAT is diluted in a suitable reaction buffer, e.g., 50 mM Hepes, pH 7.5, containing magnesium chloride or manganese chloride (1-20 mM) and a peptide or polypeptide substrate, such as myelin basic protein or casein (1-10 μ g/ml). The final concentration of the kinase is 1-20 nM. The enzyme reaction is conducted in microtiter plates to facilitate optimization of reaction conditions by

increasing assay throughput. A 96-well microtiter plate is employed using a final volume 30-100 μ l. The reaction is initiated by the addition of ^{33}P -gamma-ATP (0.5 $\mu\text{Ci/ml}$) and incubated for 0.5 to 3 hours at room temperature. Negative controls are provided by the addition of EDTA, which chelates the divalent cation (Mg^{2+} or Mn^{2+}) required for enzymatic activity. Following the incubation, the enzyme reaction is quenched using EDTA. Samples of the reaction are transferred to a 96-well glass fiber filter plate (MultiScreen, Millipore). The filters are subsequently washed with phosphate-buffered saline, dilute phosphoric acid (0.5%) or other suitable medium to remove excess radiolabeled ATP. Scintillation cocktail is added to the filter plate and the incorporated radioactivity is quantitated by scintillation counting (Wallac/Perkin Elmer). Activity is defined by the amount of radioactivity detected following subtraction of the negative control reaction value (EDTA quench).

VII. Expression analysis

All cell lines used in the following experiments are NCI (National Cancer Institute) lines, and are available from ATCC (American Type Culture Collection, Manassas, VA 20110-2209). Normal and tumor tissues are obtained from Impath, UC Davis, Clontech, Stratagene, Ardaïs, Genome Collaborative, and Ambion.

TaqMan® analysis is used to assess expression levels of the disclosed genes in various samples.

RNA is extracted from each tissue sample using Qiagen (Valencia, CA) RNeasy kits, following manufacturer's protocols, to a final concentration of 50ng/ μ l. Single stranded cDNA is then synthesized by reverse transcribing the RNA samples using random hexamers and 500ng of total RNA per reaction, following protocol 4304965 of Applied Biosystems (Foster City, CA).

Primers for expression analysis using TaqMan® assay (Applied Biosystems, Foster City, CA) are prepared according to the TaqMan® protocols, and the following criteria: a) primer pairs are designed to span introns to eliminate genomic contamination, and b) each primer pair produced only one product. Expression analysis is performed using a 7900HT instrument.

TaqMan® reactions are carried out following manufacturer's protocols, in 25 µl total volume for 96-well plates and 10 µl total volume for 384-well plates, using 300nM primer and 250 nM probe, and approximately 25ng of cDNA. The standard curve for result analysis is prepared using a universal pool of human cDNA samples, which is a mixture of cDNAs from a wide variety of tissues so that the chance that a target will be present in appreciable amounts is good. The raw data are normalized using 18S rRNA (universally expressed in all tissues and cells).

For each expression analysis, tumor tissue samples are compared with matched normal tissues from the same patient. A gene is considered overexpressed in a tumor when the level of expression of the gene is 2 fold or higher in the tumor compared with its matched normal sample. In cases where normal tissue is not available, a universal pool of cDNA samples is used instead. In these cases, a gene is considered overexpressed in a tumor sample when the difference of expression levels between a tumor sample and the average of all normal samples from the same tissue type is greater than 2 times the standard deviation of all normal samples (i.e., $\text{Tumor} - \text{average}(\text{all normal samples}) > 2 \times \text{STDEV}(\text{all normal samples})$).

A modulator identified by an assay described herein can be further validated for therapeutic effect by administration to a tumor in which the gene is overexpressed. A decrease in tumor growth confirms therapeutic utility of the modulator. Prior to treating a patient with the modulator, the likelihood that the patient will respond to treatment can be diagnosed by obtaining a tumor sample from the patient, and assaying for expression of the gene targeted by the modulator. The expression data for the gene(s) can also be used as a diagnostic marker for disease progression. The assay can be performed by expression analysis as described above, by antibody directed to the gene target, or by any other available detection method.

WHAT IS CLAIMED IS:

1. A method of identifying a candidate beta-catenin pathway modulating agent, said method comprising the steps of:
 - (a) providing an assay system comprising a MBCAT polypeptide or nucleic acid;
 - (b) contacting the assay system with a test agent under conditions whereby, but for the presence of the test agent, the system provides a reference activity; and
 - (c) detecting a test agent-biased activity of the assay system, wherein a difference between the test agent-biased activity and the reference activity identifies the test agent as a candidate beta-catenin pathway modulating agent.
2. The method of Claim 1 wherein the assay system comprises cultured cells that express the MBCAT polypeptide.
3. The method of Claim 2 wherein the cultured cells additionally have defective beta-catenin function.
4. The method of Claim 1 wherein the assay system includes a screening assay comprising a MBCAT polypeptide, and the candidate test agent is a small molecule modulator.
5. The method of Claim 4 wherein the assay is a binding assay.
6. The method of Claim 1 wherein the assay system is selected from the group consisting of an apoptosis assay system, a cell proliferation assay system, an angiogenesis assay system, and a hypoxic induction assay system.
7. The method of Claim 1 wherein the assay system includes a binding assay comprising a MBCAT polypeptide and the candidate test agent is an antibody.

8. The method of Claim 1 wherein the assay system includes an expression assay comprising a MBCAT nucleic acid and the candidate test agent is a nucleic acid modulator.
9. The method of claim 8 wherein the nucleic acid modulator is an antisense oligomer.
10. The method of Claim 8 wherein the nucleic acid modulator is a PMO.
11. The method of Claim 1 additionally comprising:
 - (d) administering the candidate beta-catenin pathway modulating agent identified in (c) to a model system comprising cells defective in beta-catenin function and, detecting a phenotypic change in the model system that indicates that the beta-catenin function is restored.
12. The method of Claim 11 wherein the model system is a mouse model with defective beta-catenin function.
13. A method for modulating a beta-catenin pathway of a cell comprising contacting a cell defective in beta-catenin function with a candidate modulator that specifically binds to a MBCAT polypeptide, whereby beta-catenin function is restored.
14. The method of claim 13 wherein the candidate modulator is administered to a vertebrate animal predetermined to have a disease or disorder resulting from a defect in beta-catenin function.
15. The method of Claim 13 wherein the candidate modulator is selected from the group consisting of an antibody and a small molecule.
16. The method of Claim 1, comprising the additional steps of:

(d) providing a secondary assay system comprising cultured cells or a non-human animal expressing MBCAT ,

(e) contacting the secondary assay system with the test agent of (b) or an agent derived therefrom under conditions whereby, but for the presence of the test agent or agent derived therefrom, the system provides a reference activity; and

(f) detecting an agent-biased activity of the second assay system,
wherein a difference between the agent-biased activity and the reference activity of the second assay system confirms the test agent or agent derived therefrom as a candidate beta-catenin pathway modulating agent,

and wherein the second assay detects an agent-biased change in the beta-catenin pathway.

17. The method of Claim 16 wherein the secondary assay system comprises cultured cells.

18. The method of Claim 16 wherein the secondary assay system comprises a non-human animal.

19. The method of Claim 18 wherein the non-human animal mis-expresses a beta-catenin pathway gene.

20. A method of modulating beta-catenin pathway in a mammalian cell comprising contacting the cell with an agent that specifically binds a MBCAT polypeptide or nucleic acid.

21. The method of Claim 20 wherein the agent is administered to a mammalian animal predetermined to have a pathology associated with the beta-catenin pathway.

22. The method of Claim 20 wherein the agent is a small molecule modulator, a nucleic acid modulator, or an antibody.

23. A method for diagnosing a disease in a patient comprising:

EX03-092P

- (a) obtaining a biological sample from the patient;
- (b) contacting the sample with a probe for MBCAT expression;
- (c) comparing results from step (b) with a control;
- (d) determining whether step (c) indicates a likelihood of disease.

24. The method of claim 23 wherein said disease is cancer.

ABSTRACT OF THE DISCLOSURE

Human MBCAT genes are identified as modulators of the beta-catenin pathway, and thus are therapeutic targets for disorders associated with defective beta-catenin function. Methods for identifying modulators of beta-catenin, comprising screening for agents that modulate the activity of MBCAT are provided.

Nucleic Acid and Polypeptide sequences

>gi|15451913|ref|NM_003933.3| Homo sapiens BAI1-associated protein 3 (BAIAP3), mRNA

ATGAGACCCCGGGGAGCAGCGTTTGCAGCGGGCCCGCCAGGTGACCTGCACCTGGGCACCGCCATCGGCT
TCGCAGGGGCCATCTGGAGGAGTCGGTCACCCGCCATGTCGACCTTGCTGGACATTAAGAGCAGCGTGCT
CAGGCAGGTGCAGGTGTGCCCGTCCCTCCGCCGAGGACTGAGCAGGACCCAGGGAGTGCCAGCGCCGAC
CCGCAGGAGCCTGCCACGGGGGCCCTGGAAACCCGGGGATGGCGTGGAGTTCCTTTGCCACATGCGCCTCA
TGCTGAAGAAGGGGAAGGCAGACAGGGCTTGCCGTGCCCTCGAGGTCCCCCTGCGCAGTGGCTCGCCAGC
ACCCCCGAGCCTGTGGATCCCAGCCTCGGCCCTGAGAGCCCTGGCCCCAGAGGAGGTGGAGATGCTCTAC
GAGGAGGCCCTGTACACGGTGCTTTACCGCGCGGGTACCATGGGCCCTGACCAGGTGGACGACGAGGAGG
CCCTGCTCAGCTATCTCCAGCAGGTGTTTGGCACAGCCTTGAGGAGCACACTGAGGCCATCGAGCGAGT
GAGGAAGGCCAAGGCCCCACGTATGCCCTGAAAGTCTCTGTTCATGCGTGCCAAGAACCTTCTGGCCAAG
GACCCCAACGGCTTCAGCGACCCATACTGCATGCTGGGCATCCTGCCTGCCTCGGACGCCACGCGGGAGC
CCCGTGCACAGAAGGAGCAGCGCTTCGGCTTCGCCAAGGGCAGCAAGCGCGGTGGACCCCTGCCTGCCAA
GTGCATCCAGGTACCCGAGGTGAAGAGCAGCACCTGAACCCCGTCTGGAAGGAGCACTTCCTCTTCGAG
AAGCGTGCAGGAAGCTGAATGAAGTCACTCGGCCTGAAGGGCATGGGCAGGTACTTCAAACAGATCGTCAA
GTACGCCCCGCGCAACCGGGACAGCAGGACCCACCGAGGACCACACCGATGACTTCTTGGGGTGCCCTCAAC
ATACCTGTCCGGGAGGTGCCTGTGGCTGGCGTGCACCGCTGGTTCAAGCTGGAGCCACGCTCCAGTGCCT
CGCGTGTGCAGGGACACTGCCACCTGGTTCTCAAGCTGATCACTACGCAGAGGGATACGGCCATGAGCCA
GCGCGGGCGATCCGGCTTCTGTCCCACCTGCTGCTGCTCAGCCATCTGCTGCGGTTGGAGCACTCAGCA
GAGGAGCCCCAATCCAGCAGCTGGCGAGGAGAGCTCAGCACACCAGCCGCCACCATCCTCTGCCTGCACG
GAGCCCAGAGCAACCTGTCAACCTTGCAGCTGGCCGTGCTGCACTGGCAGGTCAGCAGCCGCCACCATCA
AACCTGCACGCTGGACTACAGCTACCCTGCTGGGGCTGCTGGAGGACATGCAGGCACACTGGGAAGAGGCT
CCTTCACTGCCCCAGGAGCAGGAGGAGAGCCTGGCTGATAGCCTTTCCGCCTTCTCTGAGTTCCGGCTGC
AGCTGCTGCGCCAGCTCCGAGACTACTTCCCTGCCACCAACAGCACCCTGTCCACCGCTGGAGCTGCT
GCTGAAGTGTCTGGGCAAGCTGCAGCTCTTCCAACCTTCTTTGAGATCTGCCCTTCGAGTCCGGAGCTG
AACATCAGATGTCTGCGGCCCTGAAGAGAGGCCAACCCTGAGTGGTACGACAGGATCCTGAATGACAAGA
GTCCCCGAGAGCAGCCAGGACCACAGCGCTGCCTGGGCTGGTTGTGCTGGCTGACGCCGTCTATGATGA
CCTTCAGTTCTGCTACAGTGTGTACGCCAGCCTCTTCCACAGCATCCTCAATGTGGACGCTCTTACCCTG
ACCTTCCGGCAGCTGGAGCGTCTGGTGGCTGAGGAGGCGTGGGTGCTGACGGAGGAGCTGAGCCCCAAGA
TGACCTTGGAGGTGGCTCGGGCTCTTTGAGCTCTACCTGACCTGGCTGACCTCCAGCGCTTCTGGGA
TAGCATCCCTGGCCGGGACAGCCGCTCTCTGCCCTGGCTGGCATCCACGCCCCCTTCTGCCTGCTGTG
AAGCTCTGGTTCCAAGTGTCTGAGGGACCAGGCCAAGTGGAGGCTTCAGGGAGCCGTGGACATGGACACGC
TGGAGCCCGTGGACGCTCTCTCCAGGCACAGCAGCTCCGCAGCCACTGCTGGTCTCTGCCTCAGCCACAT
CCAGGAGTTGTGGGTGCGCCTGGCGTGGCTGACCTGCCAGGCTCAGGGGCTGGGCACCCAGCTTGGC
CAGGACGTGTGTGAGGCCACCTCTTCTATACGGAGCTGCTTCGGAAGAAGGTGGACACTCAGCCAGGGG
CGGCCGCTGAAGCAGTGAGCGAGGCGCTCTGCGTGGTCTCAACAATGTGGAGCTCGTGCGCAAGGCTGC
TGGGCAGGCTTGAAGGGCTTGGCATGGCCAGAGGGGGCCACGGGGCCCGAGGGGGTGCCTCCCGCCCT
CTGCTCAGCTGCACACAGGCCCTGGACGATGATCTGCAACGGGAGGCCACACGGGTGACAGCGCACCTGA
CCTCTAAGATGGTGGGCGACATCCGCAAGTATGTACAGCACATCAGTCTCTCGCTGACTCCATCCAGAA
CGATGAGGCCGTGGCCCCGCTCATGAAGTACCTGGATGAGAAGCTGGCCCTGCTGAACGCCCTCGCTGGTG
AAGGGGAACCTGAGCAGGCTGCTGGAGGCCCTGTGGGAGCTACTCTCCAGGCCATTCTGCAGGCGCTGG
GTGCAAAACCGTGACGTCTCTGCTGATTTCTACAGCCGCTTCCATTTCACGCTGGAGGCCCTGGTCAGTTT
TTTCCACGCAGAGGCTCAGGGTTTGCCCCCTGGAGAGCCTGAGGGATGGAAGCTACAAGAGGCTGAAGGAG
GAGCTGCGGCTGCACAAATGTTCCACCCGCGAGTGCATCGAGCAGTTCTACCTGGACAAGCTCAAACAGA
GGACCTTGGAGCAGAACCGGTTTGGACGCCCTGAGCGTCCGTTGCCATTACGAGGCGGCTGAGCAGCGGCT
GGCCGTGGAGGTGCTGCACGCCCGCGGACCTGCTCCCCCTGGACGCCAACGGCTTAAGTGACCCCTTTGTG
ATCGTGGAGCTGGGCCACCGCATCTCTTCCACTGGTCCGACGCCAGAGGACCCAGGTGAAGACCCGGA
CGCTGCACCTGTATACGACGAACCTTCTACTTTTCCGTGCCCTGCCGAGGCTGCCCGCCCGCGCGGCT
TGCTGTGTTGTTTCCAGCTCATGGACACGACTGGCTGTCCACCAACGACTTCTGCTGGGGAGGCGGCCCTC
GGCTAGGTGGCGTCACTGGTGTGCCCCGGCCCCAGGTGGGCGGGGTGCAAGGGCTGGGCAGCCTGTCA
CCCTGCACCTGTGCCGGCCAGAGCCAGGTGAGATCTGCGCTGAGGAGGCTGGAAGGCCGCCACAGCAA
GGAGGCGCAGGAGTTCTGTGAAGAACTCAAGGAGCTGGAGAAGTGCATGGAGGCGGACCCCTGAGTCCAT
CAGCTGCCAGCCCCGGCCCTGGCCCCACCCCAAGTTCCCTGAAGCATCCTCCAGCTCACTGTGGCCAGC
TTTGTGCAACCAGGGCCACGGCGCCCCCTCTGTGCTGTGACGTGTGTGCTGCTGGCTGGCCCCGCGGCGC

CTACCGCCCTGGCCGTGTCTGTCTGGTGTGTGCTGTGAACCCCTGCACCCAACCCACATCTGGGTGGCC
 AACTTGGCAGGACTTGGCCAGCAGCTGCCCAGGACACAGTGCAGGCCAGAGCGGGCTTGACCACCTGGTG
 GGCTTCCCTGCCCGCTTCTTGGGCTCCCCGGCCCTGGGTGGGCGGTGGGCAGCTGGTCTCCAGGGACTC
 AGTGAGTGGCTGTGCTCTCTGCACAACGGGAATGTGCAGACGCATTTTGGTAATCACAGCTGGGGAGT
 GAAAAGGGTGCCACTGGCACCCTGGGTGGATGGTCCAGAGCCTCCACCCACAGAGGGGATGCAAAGGGC
 AGGTGAGTCAAGAACCGCATAGGTCTCCAGTCCCCACGGGGCTCCCAGGCCGGGAAAGGTTCCCTGAG
 GTCACCTTGAGGCCAGGGACGTACCCAAGGCTGGTGGTCAAGTGTGAAGGGCTCCGTGCCAACTGGTCAG
 CTGTCTTTCACGCACATATCCGTGGCCACCTGAGACCTGCTCCACGACCCCTCCAGGCAGAGCCGAGAGT
 TCGCCCCAACCCCTTCCCAGGCCAGTGTGAAAAACAGACTCACAAGGGGGCTTCTTGGCCTGCAGCTTCA
 TTTGCGAGAGCGCCGAGGCAGGACACAGAGCACAGCTGTGCTGGAAGTGTGGGGAGAACCCGGACAGCTC
 AGTCTGCCAGCAGCCGAAAGAGCCGAGGCTGCCAGGCCATTTATGTCCCTCATGTCTCTAGATTTTC
 TCGTCACCCAGCCTCAAAAATATATGTGTCTGCAACCCCTCAGT

>gi|29736784|ref|XM_113950.4| Homo sapiens Munc13-4 protein
 (LOC201294), mRNA

TTCAGCCAGCCTGGGCACAGCACATGGATGGGGCAGGATGCTGCCCGCCTGTTCGGGCCCCCGCCC
 CCTGCTGAGAGCGTTGTCCCCTTCCCAGGGCCTCTGTAGGAGGGTGGGGCGGGCCAGGCCTGCCCACGG
 GGTCCACCTTCTGTCTCCAGCCAGAGGAGGGTCTCTAGGGATCCCAGCTGGGGAAGGTGCAGGGGGCCTG
 GGAGACTGCCCTCCGAAGCGGTGAAACCCCTCACCTTCCGTCCCTCCCAGCCACCCCTCCCTAAACTTC
 CCCTGACAGAGGGTGGCAGCCCCAGGCTCTTTGCATAATCCTGTGGCTTCGCTGTCTTCAACCCAGACCA
 GCGGACAGGGAAGGCGAGAGAAGGCCACCATGGCGACACTCCTCTCCCATCCGACGAGCGCCCTCCCTT
 CTTGCGCCAGGCCATCAAGATAAGGCGCCGACAGATCAGAGATCTACAGGATCCCCCGCCCCAAATGGCC
 CCGGAGATCCAGCCTCCATCCCACCCTTCTCCCCGAGCAGCGGGCCCTGCTCTACGAGGACGCACTCT
 ACACTGTCTTGCACCGCCTGGGTATCCTGAGCCCCAACCATGTGACGGAGGCCTCTGAGCTGCTGCGATA
 CCTGACAGGAGGCCCTTCCACGTGGAGCCCAGGAGCACCAGCAGACACTGCAGCGGGTCAGGGAGCTTGAG
 AAGCCAATATTTTGTCTGAAGGCAACAGTGAAACAGGCCAAGGGCATTTCTGGGCAAAGATGTGAGTGGGT
 TCAGCGACCCCTACTGCCTGCTGGGCATTGAGCAGGGGGTAGGTGTGCCAGGGGGCAGCCCCGGGTCCCG
 GCATCGGCAGAAGGCTGTGGTGAGGCACACCATCCCCGAGGAGGAGACCCACCGCAGCAGGTATCACC
 CAGACACTCAACCCCGTCTGGGACGAGACCTTCATCTGGAGTTTGAGGACATCACCATGCGAGCTTTC
 ATCTGGACATGTGGGACCTGGACACTGTGGAGTCTGTCGACAGAAGCTTGGGGAGCTCACGGATCTGCA
 TGGGCTTCGACGATCTTTAAAGAGGCCCGGAAGCAAAAGGCCAGGACGACTTTCTGGGGAACTGGTT
 CTGAGGCTGCAGGACCTGCGCTGCCGAGAGGACCAGTGGTACCCCTGGAACCCGCACTGAGACCTACC
 CAGACCGAGGCCAGTGCCACCTCCAGTTCCTCACTCATCCATAAGCGGAGAGCCACTTCGGCCAGCCGCTC
 GCAGCCGAGCTACACCGTGCACCTCCACCTCCTGCAGCAGCTTGTGTCCCACGAGGTACCCAGCAGCAG
 GCGGGAAGCACCTCTGGGACGGGTGCTGAGTCCCCAGGCTGCCACCGTCTCTTTCTGCACGCCACAC
 AGAAGGACCTATCCGACTTCCACAGTCCATGGCGCAGTGGCTGGCTTACAGCCGCCTCTACCAGAGCCT
 GGAGTTCCCCAGCAGCTGCCTCCTGCACCCCATCACCAGCATCGAGTACCAGTGGATCCAGGGTCGGCTC
 AAGGCAGAACAGCAGGAGGAGCTGGCCGCCCTCATTCAGCTCCCTGCTGACCTACGGCTCTCCCTCATCC
 GGAGGTTCCGCTCTGTCTTCCCCCTCTCTGTCTCGGACTCCCCAGCCCGGCTGCAGTCTCTTCTCAGGGT
 CCTGGTACAGATGTGCAAGATGAAGGCCCTTTGGAGAACTGTGCCCCAACACCGCCCCATTGCCCCAGCTG
 GTGACTGAGGCCCTGCAGACTGGCACCCTGAATGGTTCCACCTGAAGCAGCAGCACCATCAACCCATGG
 TGCAGGGCATCCCGGAGGCAGGCAAGGCCCTGGTGGGCTGGTACAGGATGTGATTGGCGACCTGCACCA
 GTGCCAGCGCACATGGGACAAGATCTTCCACAATAACCTCAAGATCCACCTCTTCTCCATGGCTTTCCGG
 GAGCTGCAGTGGCTGGTGGCCAAGCGGGTGCAGGACCACACGACGGTTGTGGGTGATGTAGTGTCCCCAG
 AGATGGGCGAGAGTCTGTTCAGCTCTACATCAGCCTCAAGGAGCTCTGCCAGCTGCGCATGAGCTCCTC
 AGAGAGGGATGGAGTCTTGGCCCTGGATAATTTCCACCGCTGGTTCCAGCCGGCCATCCCCTCTGGCTG
 CAGAAGACGTACAACGAGGCCCTGGCGCGGGTGCAGCGCGCTGTGCAGATGGATGAGCTGGTGCCCTGG
 GTGAAGTACCAAGCACAGCACATCAGCGGTGGATCTATCCACCTGCTTTGCCAGATCAGCCACACTGC
 CCGGCAGCTGGACTGGCCAGACCCAGAGGAGGCCCTTCATGATTACCGTCAAGTTTGTGGAGGACACCTGT
 CGCCTGGCCCTGGTGTACTGCAGCCTTATAAAGGCCCGGGCCCGCGAGCTCTCTTACAGCCAGAAGGACC
 AAGGCCAGGCAGCCAACATGCTGTGTGTGGTGGTGAATGACATGGAGCAGCTGCGGCTGGTGTATCGGCAA
 GTTGGCCCGCCAGCTGGCATGGGAGGCCCTGGAGCAGCGGGTAGGGCCGTGCTGGAGCAGGGGCAGCTG
 CAGAACACGCTGATGCCAGCTGCAGAGCGCGCTGGCCGGGCTGGGCCATGAGATCCGCACTGGCGTCC
 GCACCTTGGCCGAGCAGTGGAGGTGGGATCTGACCAAGCACATCCAGAACTGGTGGGCGTCAGGGAGTC
 TGTCTTGCCTGAGGATGCCATTCTGCCCCGATGAAGTTCTGGAGGTGGAGCTTTGTACATGAACACC
 AACTTGGTGCAGGAGAACTTCAGCAGCTCTTGACCTGCTCTGGACCCACACACTCACAGTGTGGTGG
 AGGCGGCCGCTCCCAGCGCAGCTCATCCCTGGCTTCCAACAGGCTGAAGATTGCCCTGCAGAACCTGGA

GATCTGCTTCCACGCTGAGGGCTGTGGCCTGCCACCCAAGGCCCTGCACACTGCCACCTTCCAGGCTCTG
 CAGAGGGACCTGGAGCTGCAGCGGCCCTCCAGCCGGGAACATCCGGAAGTACTTCTGCAGCCGAATCC
 AGCAGCAGGCAGAAACCACCTCTGAGGAGCTGGGGGCTGTGACAGTCAAGGCCCTCCTACCGCGCTCTGA
 GCAGAAAGCTGCGTGTGGAGCTGCTCAGCGCCTCCAGCCTGCTGCCCCTGGAATGAGCTCCAGCGAC
 CCCTTTGTCCAGCTGACCTTGGAGCCCAGGCATGAGTTCCCTGAGCTGGCCGCCCCGGGAGACCCAGAAGC
 ACAAGAAGGACCTTCACCCATTGTTTGTATGAGACCTTTGAATTCCTGGTGCCTGCTGAGCCGTGCCGCA
 GGCTGGGGCATGCCTCCTGCTCACCCTGCTGGACTACGACACGCTGGGGGCCGACGACCTGGAAGGCGAG
 GCCTTCTGCCCCGTGCGTGAGGTGCCCCGGGCTGAGTGGCTCTGAGGAGCCTGGTGAGGTGCCTCAGACCC
 GCCTGCCCCCTCAGTACCCCGCACCCAACGGGGACCCAATCCTGCAGCTGCTGGAGGGCCGGAAGGGTGA
 CCGAGAAGCCAGGTCTTTGTGAGGCTGCGGCGGCACCGGGCCAAGCAGGCCCTCCAGCATGCCTTGCGG
 CCGGCACCGTAGCCGTAGAGGTTTGCAGGTGGGGCTGGGTCCCCGGTGGGGACTTGCAAGGGCTTTCTGT
 AGGGTCTGGGGCTTCCCCGCCACATCGCGGCCCTCCAGCCTGGCCTAACACTTGGGGAGCCCCAGCATGC
 GGAGTGCCAGAGTGACAGACCTCCCTGCCTCCCATGGTATGGGGGCTCAGCAGCGACATCTCTACTCC
 CGCCTCCCTGCCTCCAGCCCTGGCTGCAATGTCTCTACCACATCCAGCACCAGGGGGAGCAAACCTGC
 CCCTGCCCCCTCTCAGAAAAGCTGCTGTGGTGGGCAGGGGATTGGGCCATCTGTCTCCTGGCCCTGGCC
 CATCTGCCCTCCTGGCCTTCTGTTCACGCCACTGGGGTGGGGGCCAGGTTCACTGGGACCAGGGCTACAG
 GCACAGAGTCTCCTGGAAGGGAGAGGGGACCTGCCAAAGATGAGGCTCCAGCTGCCCTGGGGGGAGG
 GTGGTGGCCATTACTAGAGGGGGCTGGGTCTCTCCCCAGGGGCTGCCAGCATCCAGGCCAGGAAGCCT
 GGAGCCAAAGAACCTTCTGGCTCTGAGGGAGCAAGAGCTGGCAGGCGGAGGGCTGGCACAGACAGCGGA
 AGCAGAAAGGACAGTTTGGCTGCTGTGTCTGCTGCGCACGCCCCCTCCCCGGACAGCACCTGCCACCTAG
 AAACCTTCTTAGCAAAAAAATTAATAAAAACAAATCCATTGTCTCTTAAATATGCCTTTGGCCTACAG
 TGGGGCTTGAATGCGAGCCAGGCCGGGTAGCTTCTCCCCAGCCCTCAGGGGACTTTGAGCTACCGCCA
 CCTTGGGGTAGCTACAAAGCAGGGGGGTAGGTGTGAAATAACTGAGGCAGAGGCAGGGCTAGGGTGATT
 TTTGGCCGTGGGCTTTGAATAAATTGCTTTACCAGGC

>gi|34452699|ref|NM_005845.2| Homo sapiens ATP-binding cassette, sub-
 family C (CFTR/MRP), member 4 (ABCC4), mRNA

GGACAGGCGTGGCGGCCGAGCCCCAGCATCCCTGCTTGAGGTCCAGGAGCGGAGCCCCGCGGCCACCGCC
 GCCTGATCAGCGCGACCCCGGCCCGCCCCGCCCCGCCCCGCGCAAGATGCTGCCCGTGTACCAGGAGGTGA
 AGCCCAACCCGCTGCAGGACGCGAACATCTGCTCAGCGGTGTTCTTCTGGTGGCTCAATCCCTTGTTTA
 AATTGGCCATAAACGGAGATTAGAGGAAGATGATATGTATTAGTGCTGCCAGAAGACCGCTCACAGCAC
 CTTGGAGAGGAGTTGCAAGGGTTCTGGGATAAAGAAGTTTAAAGAGCTGAGAATGACGCACAGAAGCCTT
 CTTTAAACAAGAGCAATCATAAAGTGTACTGGAATCTTATTTAGTTTGTGGGAATTTTACGTTAATTGA
 GGAAAGTGCCAAAGTAATCCAGCCCATATTTTGGGAAAAATTATTAATTATTTTGAATTTATGATCCC
 ATGGATTCTGTGGCTTTGAACACAGCGTACGCCATATGCCACGGTGTGACTTTTTGCACGCTCATTTTGG
 CTATACTGCATCACTTATATTTTATCACGTTCACTGTGCTGGGATGAGGTTACGAGTAGCCATGTGCCA
 TATGATTTATCGGAAGGCACTTCGTCTTAGTAACATGGCCATGGGGAAGACAACCACAGGCCAGATAGTC
 AATCTGCTGTCCAATGATGTGAACAAGTTTGATCAGGTGACAGTGTCTTACACTTCCTGTGGGCAGGAC
 CACTGCAGGCGATCGCAGTGAATGCCCCACTCTGGATGGAGATAGGAATATCGTGCCCTGTGTTGGGATGGC
 AGTTCTAATCATCTCCTGCCCTTGCAAAGCTGTTTGGGAAGTTGTTCTCATCACTGAGGAGTAAAACT
 GCAACTTTCACGGATGCCAGGATCAGGACCATGAATGAAGTTATAACTGGTATAAGGATAATAAAAAATGT
 ACGCCTGGGAAAAGTCATTTTCAAATCTTATTAACAAATTTGAGAAAGAAGGAGATTTCAGATTCTGAG
 AAGTTCCTGCTCAGGGGGATGAATTTGGCTTCGTTTTTCAAGTGAAGCAAAATCATCGTGTGTGTGACC
 TTCACCACCTACGTGCTCCTCGGCAGTGTGATCACAGCCAGCCGCGTGTTCGTGGCAGTGACGCTGTATG
 GGGCTGTGCGGCTGACGGTTACCTCTCTTCCCCCTCAGCCATTGAGAGGGTGTGAGAGGCAATCGTCAG
 CATCCGAAGAATCCAGACCTTTTGTCTACTTGATGAGATATCACAGCGCAACCGTCAGTGCCTGCAGAT
 GGTAAAAAGATGGTGCATGTGCAGGATTTTACTGCTTTTTGGGATAAGGCATCAGAGACCCCAACTCTAC
 AAGGCCTTTCTTTACTGTGACAGCTGGCGAATGTGTAGCTGTGGTCCGCCCCGTGGGAGCAGGGAAGTC
 ATCACTGTTAAGTCCGCTGCTCGGGGAATTTGCCCCAAGTCACGGGCTGGTCAGCGTGCATGGAAGAATT
 GCCTATGTGCTCAGCAGCCCTGGGTGTCTCGGGAACCTGAGGAGTAATATTTTATTTGGGAAGAAAT
 ATGAAAAGGAACGATATGAAAAGTCATAAAGGCTTGTGCTCTGAAAAGGATTTACAGCTGTGAGGGA
 TGGTGATCTGACTGTGATAGGAGATCGGGGAACACGCTGAGTGGAGGGCAGAAAGCACGGGTAACCTT
 GCAAGAGCAGTATCAAGATGCTGACATCTATCTCTGGACGATCCTCTCAGTGCAGTAGATGCGGAAG
 TTAGCAGACACTTGTGTTGCAACTGTGTATTTGTGCAAAATTTGATGAGAAGATCAAAATTTAGTGACTCA
 TCAGTTGCAGTACCTCAAAGCTGCAAGTCAGATTCTGATATTGAAAGATGGTAAATGGTGCAGAAGGGG
 ACTTACACTGAGTTCTTAAATCTGGTATAGATTTTGGCTCCCTTTTAAAGAAGGATAATGAGGAAAGTG
 AACAACTCCAGTTCCAGGAATCCACACTAAGGAATCGTACCTTCTCAGAGTCTTCGGTTTGGTCTCA

ACAATCTTCTAGACCCCTCCTTGAAAGATGGTGCTCTGGAGAGCCAAGATACAGAGAATGTCCCAGTTACA
 CTATCAGAGGAGAACCGTTCTGAAGGAAAAAGTTGGTTTTTCAGGCCTATAAGAATTACTTTCAGAGCTGGTG
 CTCACTGGATTGTCTTCATTTTCTTATTTCTCTAAACACTGCAGCTCAGGTTGCCATGTGCTTCAAGA
 TTGGTGGCTTTCATACTGGGCAACAAACAAAGTATGCTAAATGTCACTGTAAATGGAGGAGGAAATGTA
 ACCGAGAAGCTAGATCTTAACGGTACTTAGGAATTTATTCAGGTTTAACTGTAGCTACCGTTCTTTTTTG
 GCATAGCAAGATCTCTATTGGTATTCTACGTCTTGTAACTCTTCACAAACTTTGCACAACAAAATGTT
 TGAGTCAATTTCTGAAAGCTCCGGTATTATTTCTTTGATAGAAATCCAATAGGAAGAATTTTAAATCGTTTC
 TCCAAAGACATTGGACACTTGGATGATTTGCTGCGCTGACGTTTTTAGATTTTCATCCAGACATTGCTAC
 AAGTGGTTGGTGTGGTCTCTGTGGCTGTGGCCGTGATTCCCTTGGATCGCAATACCCCTGGTTCCCTTGG
 AATCATTTTTCATTTTCTTCGGCGATATTTTTTGGAAACGTCAAGAGATGTGAAGCGCTGGAATCTACA
 ACTCGGAGTCCAGTGTTTTCCACTTGTCTATCTTCTTCCAGGGCTCTGGACCATCCGGGCATACAAAG
 CAGAAGAGAGGTGTGAGAACTGTTTGTATGCACACCAGGATTACATTTCAGAGGCTTGGTTCTTGTTTTT
 GACAACGTCCCGCTGGTTCGCCGTCCGTCTGGATGCCATCTGTGCCATGTTTGTCTATCATCGTTGCTTT
 GGGTCCCTGATTCTGGCAAAACCTCTGGATGCCGGGCAGGTTGGTTTGGCACTGTCTATGCCCTCACGC
 TCATGGGGATGTTTCAGTGGTGTGTTTCGACAAAAGTGTGAAGTTGAGAATATGATGATCTCAGTAGAAAG
 GGTCAATTGAATACACAGACCTTGAAAAGAAGCACCTTGGGAATATCAGAAACGCCCACCAGCCTGG
 CCCCATGAAGGAGTGATAATCTTTGACAATGTGAATTCATGTACAGTCCAGGTGGGCTCTGGTACTGA
 AGCATCTGACAGCACTATTAAATCACAAGAAAAGTTGGCATTTGTGGGAAGAACCGAGCTGGAAAG
 TCCCTCATCTCAGCCCTTTTGTAGATTGTGCAAGCCGAAGGTAAAATTTGGATTGATAAGATCTTGACA
 ACTGAAATTTGGACTTCACGATTTAAGGAAGAAAATGTCAATCATACCTCAGGAACCTGTTTTGTTCACTG
 GAACAATGAGGAAAAACCTGGATCCCTTTAAGGAGCACACGGATGAGGAACTGTGGAATGCCTTACAAGA
 GGTACAACCTTAAAGAAACCATTTGAAGATCTTCTGGTAAAATGGATACTGAATTAGCAGAATCAGGATCC
 AATTTTAGTGTGGACAAAGACAACGGTGTGCTTGGCAGGGCAATTCTCAGGAAAAATCAGATATTGA
 TTATTGATGAAGCGACGGCAATGTGGATCCAAGAACTGATGAGTTAATACAAAAAAAATCCGGGAGAA
 ATTTGCCCACTGCACCGTGTGAACCATTTGCACACAGATTGAACACCATTATTGACAGCGACAAGATAATG
 GTTTTAGATTTCAGGAAGACTGAAAGAATATGATGAGCCGTATGTTTTGCTGCAAAATAAAGAGAGCCTAT
 TTTACAAGATGGTGCAACAACCTGGGCAAGGCAGAAGCCGTGCCCTCACTGAAACAGCAAAACAGGTATA
 CTTCAAAAGAAATATCCACATATTTGGTCACACTGACCACATGGTTACAAACACTTCCAATGGACAGCCC
 TCGACCTTAACATATTTTCGAGACAGCACTGTGAATCCAACCAAAATGTCAAGTCCGTTCGAAGGCATTT
 TCCACTAGTTTTTGGACTATGTAAACCACATTTGTACTTTTTTTTACTTTTGGCAACAAATATTTATACATA
 CAAGATGCTAGTTTCATTTGAATATTTCTCCAACTTATCCAAAGGATCTCCAGCTCTAACAAAAATGGTTTA
 TTTTTATTTAAATGTCAATAGTTGTTTTTTTAAAAATCCAAATCAGAGGTGCAGGCCACCAGTTAAATGCCG
 TCTATCAGGTTTTGTGCCCTTAAGAGACTACAGAGTCAAAGCTCATTTTTTAAAGGAGTAGGACAAAGTTGT
 CACAGGTTTTTGTGTTGTTTTTATTGCCCCCAAATACATGTTAATTTCCATTTATATCAGGGATTCT
 ATTTACTTGAAGACTGTGAAGTTGCCATTTTGTCTCATTGTTTTCTTTGACATAACTAGGATCCATTATT
 TCCCTGAAGGCTTCTTGTGTAGAAAATAGTACAGTTACAACCAATAGGAACAACAAAAAGAAAAGTTTG
 TGACATTGTAGTAGGGAGTGTGTACCCCTTACTCCCCATCAAAAAAAAAAATGGATACATGGTTAAAGGA
 TAGAAGGGCAATATTTTATCATATGTTCTAAAAAGAGAAGGAAGAGAAAATACTACTTTCTCAAAATGGAA
 GCCCTTAAAGGTGCTTTGATACTGAAGGACACAAATGTGACCGTCCATCTCTTTAGAGTTGCATGACT
 TGGACACGGTAACTGTTGCAGTTTTAGACTCAGCATTGTGACACTTCCAAGAAGGCCAAACCTCTAACC
 GACATTCTGAAATACGTGGCATTATTTCTTTTGGATTCTCATTTATGGAAGGCTAACCTCTGTTGA
 CTGTAAGCCTTTTTGCTTTGGGCTGTATTGAAATCTTTCTAAATTGCATGAATAGGCTCTGCTAACGTGA
 TGAGACAAACTGAAAATTTATTGCAAGCATTGACTATAATTATGCAGTACGTTCTCAGGATGCATCCAGGG
 GTTCATTTTCATGAGCCTGTCCAGGTTAGTTTACTCTGACCACATAAGCATTTGTCTATTGGGCTTTCT
 GTTGAATGAATCAACAAACCACAATACCTCTGGGACCTTTTGTACTTTATTTGAACTATGAGTCTTTAA
 TTTTCTCTGATGATGGTGGCTGTAATATGTTGAGTTTCACTTAAAGGTTTTACTATTATGGTTTGAA
 GTGGAGTCTCATGACCTCTCAGAATAAGGTGTACCTCCCTGAAATGTCATATATGTATATAGACATGCA
 CACGTGTGCATTTGTTTGTATACATATATTTGTCTTCGTATAGCAAGTTTTTGTCTATCAGCAGAGAG
 CAACAGATGTTTTATTGAGTGAAGCCTTAAAAAGCACACACCACACAGCTAACTGCCAAAATACATTG
 ACCGTAGTAGCTGTTCAACTCTAGTACTTAGAAAATACACGTATGGTTAATGTTTCAGTCCAACAAACCAC
 ACACAGTAAATGTTTATTAATAGTCATGGTTCGTATTTTAGGTGACTGAAATTGCATCAGTGATCATAAT
 GAGGTTTGTAAAACGATAGCTATATTCAAAATGTCTATATGTTTATTTGGACTTTTGAGGTTAAAGACA
 GTCATATAAACGTCCTGTTCTGTTTTTAATGTTATCATAGAATTTTTTAATGAAACTAAAATCAATTGAA
 ATAAATGATAGTTTTTCATCTCCAAAAA

>gi|27479422|ref|XM_058661.6| Homo sapiens chromosome 14 open reading
frame 35 (C14orf35), mRNA

AAACCTTTACCCGGCCTAACAGCTGAGGCGCTTTACGGCGACGGCGGCTGAGTGAGAACCTTTGGCGGCCTGT
 GGAGGCTGCCGCGGCTGCGAAGGAGGCGCGCGGCTGGCGGAGGAAGAGGAGTGGCGGCAGCGGCGGCGG
 GGACCCGTGCGGGGATGGCGGAGGTACCGCCTGGGCTTAGCAGCCTCCTCCCACCACCAGCACCTCCGGC
 CCCGGCGGCGGTCGAGCCCCGCTGTCCCTTCCCGCGGGGGCCGCCCTCGCCTGCTGCAGCGAGGACGAG
 GAGGACGACGAAGAGCACGAAGGCGGCGGCGAGCAGGAGCCCGCGGGCGGAGAGTCCGGCGACGGTGGCGG
 CCAAGGGGCATCCGTGCCTCCGCTGCCCTCAGCCGCCGAGGAGCAGCAGCAGCTCAACGGATTGATTAG
 CCCCAGACTGCCGCACCTCCGGGCGGCCGCTCCCTCAAGAGCAAGGTCCTGAGCGTAGCAGAGGTGGCC
 GCGACACAGCCACCCCTGACGGAGGCCAGAGCGACTGCAACAAAAGGAGCCGGGGTACACTCGGGCG
 AGAGGCCCCCTACTCCCTCTCTAGTAATGCAAGAAGTCCGCTCCCCAGCCCGGTGGAGGCAGCGGCGG
 GAGCGATCCCGCGGCGGCCCGCAATGGACTGGCCGAGGGCACCGAGCAGGAGGAGGAGGAGGAAGACGAG
 CAGGTGCGGCTGCTGTCTTCTGCTCCCTGACCGCCGACTGCAGCTTAAGAAGCCCTTCGGGCAGGGAGGTTG
 AGCCTGGGGAGGATCGGACGATACGATATGTCCGATATGAATCCGAGCTACAAATGCCCGATATCATGAG
 ACTGATCACCAAAGATCTGTCCGAACCTACTCCATTTATACCTATAGATATTTTATCCACAAGTGGCCA
 CAGCTGTGCTTCTTGGCCATGGTAGGGGAGGAGTGTGTAGGTGCCATCGTTTGCAAGTTGATATGCACA
 AAAAGATGTTCCGCGAGAGGTTATATAGCCATGTTAGCCGTGGATTCCAAATACAGGAGAAATGGCATTGG
 TACTAACTTGGTTAAGAAAGCTATATATGCCATGGTTGAGGGAGACTGTGATGAGGTTGTTTGGAAACC
 GAAATAACAAATAAGTCCGCTTTGAACTTTATGAAAATCTTGGTTTGTTCGAGATAAGAGGCTGTTCA
 GATACTATTTAAATGGAGTTGATGCACTGCGACTTAACTGTGGCTGCGTTGAGAACTGACATCAAGGA
 ACAACTATCATCCGCACAGAATCGACCTTTGCATGCAATGCAATTTGTACAGAATTGCTTTTCAGGTGGA
 TTTAGTAATTTCCATGCGACTCTTACCCTGTCACTGTCTCATTTGAGTGTGCGACAATATTTGTTGCACTTT
 GGCATGGCACATTTGTTCTGAATTAAGATTTGTTTAACTTCAGGAGTTCTTTTGGTACCAACAAGAT
 GTGCCAGTTGATAGCCAAGATTTATGTGTTTCATTTGCAAAGTCTGCTGACAATGTTATTTACAGTGAT
 CATTTTATCACAGAACCAGTAAGTGAACATAATTTTGTTCCTTAAAAAGCCAATGTGGATTGTAAAA
 GTCTTTAAGTATACTAACATTTACACAAAACCTGCCCTAGTTTCTGAAGTGGGTGAGGAGACGCTTC
 AGTTTTAGGTTTATTTTTCATATTAATTTTCCATTCTTGAATATTGGTACCTCAGTGATTAGTGAA
 TGAAAAAATGTAGGGTGGGTATGTCTTACAAATGAGTAAAGGTAACAATTAATTTTGTCTGCCAGTGCC
 TGTGTAGATAAGTATATTTGTCTTCATCTCTAGTTTGTGAATGCATGCTATCTTTTCTTTTCTTTAAGG
 CCTTGTGAAGCAAACCTTTGTTTATTTAAATTTCTAAATTTGATAAATATTTTCAGATTTTATAATTT
 GGATACTTTTTTCAGGTGAATGAAAGAATGGTTTACTTTAGAAGTCCCTTTTTCTTACAGTAACAAGTT
 GAATCTACTTGGAAAATTTGAGAAATGGCTCAAAAGAGATAAGAAAAGTTGATGGAGCCGGGAATGCTGG
 GGTTAGATGCACTTTTCTTTTGGAGTAAGGGAAGTTTGGAAAAGAATAGAAAATTAGTGTAAGTTG
 ATATGATTTTTATTTAATCAAAATTACTGCTACGCTGCGAAGAACAGCTTTTACAAAGTAGCTGAATTTGT
 TTTTCCCACTTGATTTGGATTACATTTGCTTTTCAATTTCTTAAATGCTTCACTTCAGGTTCTTGGTCTTG
 GAAATAAATTTCAAGGTGCATTGTATCCATTTTAAGCTGCTTTATTTTATTTTCACTTGTATGAGCAAT
 TCTTGGGGGAGCTTTGCTTTTCTTCTGCCAGAAAAACAAAAGGGGGAAATGAAAATCTTTTTTGAATGA
 GTTCTGTGGGTTTTCTTAAACAGCCACCATGTTTATTAGTTACATTGTGTTTTGGCCAATCAGTGCAATGT
 AACAAATTTTACAGTTAATGCTTTCAATTGAGTCAGTAAACCTGTGATAGATAATTTATTTAACTGGAA
 AACCTAGGTACCCATAAGAAAAAGATTCAATCTCTGTGAAAACCTGAGGAATCTGTTGTTGTTTTCATT
 TGAATATGCTCTACTTCTGCTCTAGTATTTGGTTTGGAAATATATTTTGTGGCTCTAATTACTGTATTTT
 AAAAACCTTACCTCCATTAACAGTTGGTAAAGGCCCTTTTCAGGAAAGTTTGTGCTTTTTTTTTTTTT
 TTAAGGAAAGCTGCTCTTGTCTCAGTATAGTGTTTGAAGTGAACATAGTAACAAATACTTAAAAAT
 AAAGATACACAATTTATATTTG

>gi|7705623|ref|NM_016063.1| Homo sapiens CGI-130 protein (CGI-130), mRNA

GATTCGGCACTAGCGGGGAGGAGCTTCCCGCGGCTGCTCCGCCAGCCGGGGTGGTGGCCGCATGGCTT
 CGGTCTCCTCTGCGACCTTCTCGGGCCACGGGGCTCGGTCCCTACTGCAGTTCTTGGCGCTGGTAGGGCA
 GCTCAAGAGAGTCCCACGAAGTGGCTGGGTATACAGAAATGTCCAGAGGCCGGAGAGCGTTTCAGATCAC
 ATGTACCGGATGGCAGTTATGGCTATGGTGATCAAGATGACCGTCTTAACAAAGACCCGGAAGCTATGA
 AGCAGATAACCCAGCTCCTACCAGAGGACCTCAGAAAGGAGCTCTATGAACCTTTGGGAAGAGTACGAGAC
 CCAATCTAGTGCAGAAGCCAAATTTGTGAAGCAGCTAGACCAATGTGAAATGATTCTTCAAGCATCTGAA
 TATGAAGACCTTGAACACAAACCTGGGAGACTGCAAGACTTCTATGATTCCACAGCAGGAAAAATCAATC
 ACCCTGAGATAGTCCAGCTTGTCTTGAACCTTGAGGCAGAAAGAAGCACTAACATAGCTGCAGCTGCCAG
 TGAGCCACACTCTTGAGACACTCTTAAATTTGCTGCACTCTGTAAACAAACATTATTTTCCATTTCATT
 GTATTGTGTTTTGCCATTGTGTTGGTCTGTTGATTTCCCTAGATGTGAGTCTGTTTGTGTTTCAATTGTCTGA
 ACTTCAGCAAGAAATGTGATACAACCTTGGGCACATAAAGAAGCCACAGAACAGGAAGCGGTATGAAAGT
 GCCATGGATGAACACTGGAGGTGGCAGTGCCCTGTTTATGAACATAAATAAATAAATATTAACACCTAAAA

TATTAGAATATTTATTGGAGATTTAAATCATCTTATTCTGACTTAATTACCGATATCCCCGAAGGCTAG
 GTTCATTGAATAATAGAAAATTTTCATTATGATTGCTTTTAAAGAACAGATTCCTCAGCTGATTTAGTGATA
 AGAATCCAGAAAAGAAAATGTACTAGTGATGATTCTCTCCCCAGATGAAATTGCTGCCTTATTTCAGATT
 TACTCTCTTGAGCCAGATTTTGAATTTCACTGCAGACTGCTTCAGACTTCTAATCATAGGCTTGTAACC
 TACTAATAGGCTCTGCCCCCTCTTCCCAATACTTTTTGTCACTTTAGAGATATAAACCGGGGCATATAAAAA
 TGCAACTTGTATTCCTTTGTATATTTTCCCTGTCTGACTTATAAATCTTGAGACCTTTATTGTAAAAGC
 ATTTATCATCAGGTGAGAAATATAAATAGGAACCTGGGGTCATTGAGCCTCAGGTAGGGAATATATCAACC
 CGATTTCTTCTCTCTTTTCCCTTTTATAGGATAAATAATCC

>gi|22202621|ref|NM_001513.2| Homo sapiens glutathione transferase zeta
 1 (maleylacetoacetate isomerase) (GSTZ1), transcript variant 3, mRNA
 CCGGGAGGGTTGGGCCAGAGGAGCGGTCCTGGTAGGGAGTGCCTAGCAGGGTGGAGTCGCTGTCCGGTC
 GCGCTTCACTTCTGCTCCGCCCTCCCTGCTCCCCATAACAGACCCCTCCCTCCACTAAGGCTTCCAAACG
 TCGCCCCAAGCCTCCCAATTTCTCGCAGCCCTTCATGCTCTCTCTTGGCCCTAGCAGGAGTGGTCACCA
 GCAGTCCCTGTCCCTTGGCACCAGGAGTCCCTGTAGCCTGGCTAGGTAAGGAAAGGGAGTTGCATCTCCA
 AATAAGGAAGATCCAGTTGTGGCCACCCAATTCCTCTCTTTCGAAAGACACTTTCTGCCCATCACCTGCT
 TTCTGTGCCAGGACCATCTTTACTTTCTCCACAAAAGCCCATCTCTATTCTTATTTCCGAAGCTCTGCT
 TCATGGAGAGTTCGAATTGCTCTGGCCCTGAAAGGCATCGACTACGAGACGGTGGCCATCAATCTCATAA
 AGGATGGGGGCCAACAGTTTCTTAAGGACTTCCAGGCATGAAATCCTATGAAGCAGGTGCCAACCTCGAA
 GATTGATGGAATCACCATTACCAAGTCACTGGCCATCATGAGTATCTAGAGGAGATGCGTCCCACTCCG
 CGACTTCTGCCCTCAGGACCCAAAGAAGAGGGCCAGCGTGCCTATGATTTCTGACCTCATCGCTGGTGGCA
 TCCAGCCCCCTGCAGAACCTGTCTGTCTGAAAGCAAGTGGGAGAGGAGATGCAGCTGACCTGGGCCCAGAA
 CGCCATCACTTGTGGCTTTAACGCCCTGGAGCAGATCCTACAGAGCACAGCGGCATATACTGTGTAGGA
 GACGAGGTGACCATGGCTGATCTGTGCTTGGTGCCTCAGGTGGCAAATGCTGAAAGATTCAAGGTGGATC
 TCACCCCTTACCCTACCATCAGCTCCATCAACAAGAGGCTGCTGGTCTTGGAGGCCCTCCAGGTGTCTCA
 CCCCTGCCGGCAGCCAGATACACCCACTGAGCTGAGGGCCTAGCTCCCAAATCCTGCCCCGTTGGCACAG
 GGCCACAGGAGCAGAGCTGGGTGGGCTGAAGAGGCCTGGAAACGAGAGTCTTAATTGAGGAGATGGGAG
 ACTCGAACTCTAGCCCTGGATCTGCCCTTCTGCTGAAACTTGTTCACCTCAGTCCCCCTCATCTGTACAC
 CGCATGTGGGGTGGAGTAGGGAGATGCGGGGAGCAGGGTGGGCAGGAATACTGTTATCTATGTGACGGGG
 CAGTCGTGAGGCTGAGATGAGAATGCGGATTAATGCTGGCGTGTCTACCGTAACACCACGGGGAAGG
 CTGTGTGCCTTTTCTCATCCGCTTTTGTGTGTGTGACTCCAAAGAATGCCCGCGCTGAAATTTGGCGTG
 AATTAAACTGAAGCCAGGCCTCTAAAAAAAAAAAAAAAAAAAAA

>gi|21361432|ref|NM_014264.2| Homo sapiens serine/threonine kinase 18
 (STK18), mRNA
 CAGAGGGCACCGCCAGGCTCGGAAGGTGTCAGGGAGAACTTCCGTTGGTTTCAGCGTCGTCGCCTGGA
 GCGGGCGTTTAGAGAGCCGAGCCTGATGGGCCTCAAGGCCGGCTGGCTGCTTGGAGCGCTGCCTCGAAGG
 GACTGCGTAAGGAAGCTAATCCGGAGAACCAGGCCAGAGCCTGAAATATGGCGACCTGCATCGGGGAGA
 AGATCGAGGATTTTAAAGTTGGAATCTGCTTGGTAAAGGATCATTTGCTGGTGTCTACAGAGCTGAGTC
 CATTCACAGTGGTTTGAAGTTGCAATCAAAATGATAGATAAGAAAGCCATGTACAAAGCAGGAATGGTA
 CAGAGAGTCCAAAATGAGGTGAAAATACATTGCCAATTGAAACATCCTTCTATCTTGGAGCTTTATAACT
 ATTTTGAAGATAGCAATTATGTGTATCTGGTATTAGAAATGTGCCATAATGGAGAAATGAACAGGTATCT
 AAAGAATAGAGTGAACCCCTTCTCAGAAAATGAAGCTCGACACTTCATGCACCAGATCATCACAGGGATG
 TTGTATCTTTCATTCTCATGGTATACTACACCGGACCTCACACTTTCTAACCTCCTACTGACTCGTAATA
 TGAACATCAAGATTGCTGATTTTGGGCTGGCAACTCAACTGAAAATGCCACATGAAAAGCACTATACATT
 ATGTGGAACCTCTAACTACATTTACCCAGAAATTGCCACTCGAAGTGCACATGGCCTTGAATCTGATGTT
 TGGTCCCTGGGCTGTATGTTTTATACATTACTTATCGGGAGACCACCTTTCGACACTGACACAGTCAAGA
 ACACATTAAATAAAGTAGTATTGGCAGATTATGAAATGCCAATTTTTTTGTCAATAGAGGCCAAGGACCT
 TATTCACCAGTTACTTCGTAGAAATCCAGCAGATCGTTTAAAGTCTGTCTTCAGTATTGGACCATCCTTTT
 ATGTCCCCGAAATCTTCAACAAAAAGTAAAGATTTAGGAACGTGGAAGACTCAATTGATAGTGGGCATG
 CCACAATTTCTACTGCAATTACAGCTTCTTCCAGTACCAGTATAAGTGGTAGTTTATTTGACAAAAGAAG
 ACTTTTGAATTGGTCAGCCACTCCCAAATAAAATGACTGTATTTCCAAAGAATAAAAGTTCAACTGATTTT
 TCTTCTTCAGGAGATGGAACAGTTTTTATACTCAGTGGGGAATCAAGAAACCAGTAATAGTGAAGGG
 GAAGAGTAATTCAGATGCAGAAAGAGGCCACATCTCGATACCTTCGTAGAGCTTATTTCTCTGTATAG
 ATCTGGCACTTCTCAAGTAGACAGTCTCAAGCAAAAACATATACAATGGAACGATGTCACTCAGCAGAAATG
 CTTTCAGTGTCCAAAAGATCAGGAGGAGGTGAAAATGAAGAGAGGTACTCACCCACAGACAACAATGCCA
 ACATTTTAACTTCTTTTAAAGAAAAGACATCCAGTAGTTCTGGATCTTTTGAAGACCTGATAACAATCA

AGCACTCTCCAATCATCTTTGTCCAGGAAAACTCCTTTTCCATTTGCAGACCCGACACCTCAGACTGAA
 ACCGTACAACAGTGGTTTTGGGAATCTGCAAATAAATGCTCATTTAAGAAAACTACTGAATATGACAGCA
 TCAGCCCAAAACCGGGACTTCCAGGGCCATCCAGATTTGCAGAAGGACACATCAAAAAATGCCTGGACTGA
 TACAAAAGTCAAAAAGAACTCTGATGCTTCTGATAATGCACATTCTGTAAAACAGCAAAATACCATGAAA
 TATATGACTGCACCTTCACAGTAAACCTGAGATAATCCAACAAGAATGTGTTTTTGGCTCAGATCCTCTTT
 CTGAACAGAGCAAGACTAGGGGTATGGAGCCACCATGGGGTTATCAGAATCGTACATTAAGAAGCATTAC
 ATCTCCGTTGGTTGCTCACAGGTTAAACCAATCAGACAGAAAACCAAAAAGGCTGTGGTGAGCATACTT
 GATTACAGAGGGTGTGTGTGGAGCTTGTAAAGGAGTATGCATCTCAAGAATATGTGAAAGAAGTTCCTC
 AGATATCTAGTGATGGAAATACGATCACTATTTATTATCCAAATGGTGGTAGAGGTTTTCTCTTGCTGA
 TAGACCACCTCACCTACTGACAACATCAGTAGGTACAGCTTTGACAATTTACCAGAAAAATACTGGCGA
 AAATATCAATATGCTTCCAGGTTTGTACAGCTTCTAAGATCTAAATCTCCCAAAATCACTTATTTTACAA
 GATATGCTAAATGCATTTTGTATGGAGAATTCTCCTGGTGTGATTTTGAAGTTTGGTTTTATGATGGGGT
 AAAAATACACAAAACAGAAGATTTTCAATCAGGTGATTGAAAAGACAGGGAAGTCTTACACTTTAAAAAGT
 GAAAGTGAAGTTAATAGCTTGAAAGAGGAGATAAAAATGTATATGGACCATGCTAATGAGGGTCATCGTA
 TTTGTTTAGCACTGGAATCCATAATTTTCAAGAAGAGGAAAGGAAAACTAGGAGTGCCTCCCTTTTCCCAAT
 AATCATAGGAAGAAAACTGGTAGTACTAGTTACCTAAGGCCTTATCACCTCCTCCTTCTGTGGATTCA
 AATTACCCAACGAGAGATAGAGCATCTTCAACAGAATGGTCATGCATAGTGATGCTTCTCCAACACAGG
 CACCAATCCTTAATCCCTCTATGGTTACAAATGAAGGACTTGGTCTTACAACACAGCTTCTGGAACAGA
 CATCTCTTCTAATAGTCTAAAAGATTGTCTTCTTAAATCAGCACAACTTTGAATCTGTTTTTGTGAAA
 AATGTTGGTTGGGCTACACAGTTAACTAGTGGAGCTGTGTGGGTTTCAAGTTAATGATGGGTCCCAGTTGG
 TTGTGCAGGCAGGAGTGTCTTCTATCAGTTATACCTCACCAATGGTCAAAACAACCTAGGTATGGAGAAAA
 TGAAAAATTACCAGACTACATCAAAACAGAAATTACAGTGTCTGTCTTCCATCCTTTTGATGTTTTCTAAT
 CCGACTCCTAATTTTTCATTGATTAAACTCCTTTTACAGACATATAAGTTTAATAAATAACTTTTTTGTGGA
 CTTTCAAGTAAAGTGATTTTTTTTTTAATTTAACATAAAGTCTTCAGAAAGCCTTTCATGAAAGAATTTTA
 ACCTATAATGTAACCATGTATCTGAGATAACAAAGCAGAAATGAACTTGAGTCACCTACTAAATATAGT
 GGATATAAAATAGAACACCTGACTTTGCTCTTAGACCATAA

>gi|37552193|ref|XM_166453.4| Homo sapiens tau tubulin kinase 1
 (TTBK1), mRNA

ATGGCCCATCATGAATTAGAAGGTGGTTTCCACACCTCCACGTCCACCCCTGTCCCCTATGGCAGGATGT
 CTAGATGCAAGAGCCCTGGGAAGTTATTTCAATCCCAAGGAAAGGGCAGGCCCGGCGACACCCCTCCCTC
 TGGCTGGCGGATGCAGTGCCTAGCGGCCGCCCTTAAGGACGAAACCAACATGAGTGGGGGAGGGGAGCAG
 GCCGACATCCTGCCGGCCAACCTACGTGGTCAAGGATCGCTGGAAGGTGCTGAAAAAGATCGGGGGCGGGG
 GCTTTGGTGAGATCTACGAGGCCATGGACCTGTGACCAGGGAGAATGTGGCCCTCAAGGTGGAGTCAGC
 CCAGCAGCCCAAGCAGGTCTCAAGATGGAGGTGGCCGTGCTCAAGAAGTTGCAAGGGAAGGACCATGTG
 TGCAGGTTTATTGGCTGTGGCAGGAACGAGAAGTTAACTATGTAGTGATGCAGCTCCAGGGCCGGAACC
 TGGCCGACCTGCGCCGTAGCCAGCCGCGAGGCACCTTCACGTGAGCACCACATTGCGGCTGGGCAAGCA
 GATCTTGGAGTCCATCGAGGCCATCCACTCTGTGGGCTTCTTCACCGTGACATCAAGCCTTCAAACCTTT
 GCCATGGGCAGGCTGCCCTCCACCTACAGGAAGTGCTATATGCTGGACTTCGGGCTGGCCCCGCGAGTACA
 CCAACACCACGGGGGATGTGCGGCCCTTCGGAATGTGGCCGGGTTTCGAGGAACGGTTCGCTATGCCTC
 AGTCAATGCCACAAGAACCAGGAGATGGGCCGCCACGACGACCTGTGGTCCCTCTTCTACATGCTGGTG
 GAGTTTGCAGTGGGCCAGCTGCCCTGGAGGAAGATCAAGGACAAGGAACAGGTAGGGATGATCAAGGAGA
 AGTATGAGCACCAGATGCTGCTGAAGCAGATGCCGTGAGGTTCCACCTCTTCTGGACCACATTGCCAG
 CCTCGACTACTTCAACCAAGCCGACTACCAAGTTGATCATGTGAGTGTGAGAACAGCATGAAGGAGAGG
 GGCATTGCCGAGAATGAGGCCTTTGACTGGGAGAAGGCAGGCACCGATGCCCTCCTGTCCACGAGCACCT
 CTACCCCGCCCCAGCAGAACACCCGGCAGACGGCAGCCATGTTTGGGGTGGTCAATGTGACGCCAGTGCC
 TGGGGACCTGCTCCGGGAGAACACCGAGGATGTGCTACAGGGAGAGCACCTGAGTGACCAGGAGAATGCA
 CCCCCAATTTGCCCGGGAGGCCCTCTGAGGGGCTGGGCCCCAGTCCCCACCTTGTCCCCACCCCGGGG
 GTCCTGAGGCTGAAGTCTGGGAGGAGACAGATGTCAACCGGAACAACTCCGGATCAACATCGGCAAAAG
 CCCCTGTGTGGAGGAGGAACAGAGCCGAGGCATGGGGGTCCCCAGCTCCCCAGTGCCTGCCCCCAGAC
 TCCCCACAACCCAGTCCGTTCTCTGCGCTACCGGAGGGTGAACAGCCCTGAGTCAGAAAGGCTGTCCA
 CGGCGGACGGGCGAGTGGAGCTACCTGAGAGGAGGTACGGATGGATCTGCTTGGCTCGCCCTCGCGCCA
 GGCCTGCTCTCTCAGCCAGCCAGATGTGTGAGTGACACAGGCCACGCTGACCGACAGGCCAGTGCC
 CGCATGGAGCTGTGACCTCTGTGGAGCAGGAGGCCCTGAGCAACGCCCTTCCGCTCGGTGCGGCTGGCTG
 AGGAGGAGGATTTTCAGACGAAAGAGTGGGTATCATCTGACAAAGGAGACGGAGCTCAAGGACTTCCCTCC
 AGGGGCTGAGCCAGCACATCGGGCACCAAGGATGAGGAGCCCGAGGAGCTGCGGCCACTGCCCGAGGAG
 GGCAAGAGCGGCGGCGGCTGGGGGCAGAGCCACCGTCCGGCCCCGGGACGCAGCATGCAGGCGCTGG

CGGAGGAGGACCTGCAGCATTTGCCGCCCCAGCCCCTGCCACCCCAGCTGAGCCAGGGCGATGGCCGTTT
 CGAGACGTCACAGCCCCACGCCCTGGCAGCCCTTCCACTCACCCTGCACTCGGGACCCCGCCCTCGA
 CGGAGAGAGTCCGACCCACAGGCCACAGAGACAGGTGTTCTCCGTGGCGCCCCATTTGAGGTGAATG
 GCCTCCACGAGCTGTGCCTCTGAGCTGCCCTACCAGGACTTCAAAAGAGACCTCTCCGATTACCGAGA
 ACGGGCGCGGTGCTCAACAGGGTCCGGAGGGTGGGCTTCTCGCACATGCTGCTCACCACCCCCAGGTC
 CCACTGGCTCCTGTTACGCTCAGGCTAATGGGAAGGAGGAAGAGGAGGAGGAGGAGGAAGATGAGGAAG
 AGGAAGAAGAGGATGAGGAAGAAGAAGAGGAGGAAGAGGAAGAGGAGGAGGAAGAAGAGGAGGAGGAGGA
 AGAGGAGGAGGAGGCTGCAGCGGCACTTGCCTTGGGGGAGGTGCTGGGGCCTCGTAGTGCTCCAGCAGT
 GAGGGGAGTGAGAGGAGCACTGACCGGAGCCAGGAGGGTGGCCCGTCCACGCTGCTGGCAGACGATCAGA
 AGGAGTCCAGGGGCGGGCCTCCATGGCCGATGGGGACCTGGAGCCTGAGGAGGGCTCCAAAACGCTGGT
 GCTTGCTCTCTCTGGCGACATGAAGAAGTCGCCCCGTCAGTGGCGAACTGGCCCCGACCCCGACCTGGGC
 ACCCTGGCTGCCCTCACTCCTCAGCATGAGCGCCCCAGCCACGGGCAGCCAGCTGGACGTATCTGAGC
 CAGGCACCCTGTCTCTGTCTCAAGTCTGAGCCCAAGCCCCGGGGCCTGGGGCAGGGCTGGGGGCGCG
 GACAGTGACCACAGGGGTGGGGGCGTGGCAGTCACCTCCTCACCCTTCAACAAAGTTGAGAGGACCTTT
 GTGCACATTTGCGGAGAAAACCCACCTCAACGTCATGTCTTCCGGTGGACAAGCCTTGCGGTCTGAGGAGT
 TCAGCGCTGGGGGCGAGCTGGGTCTGGAGCTGGCCTCTGATGGGGGCGCTGTGGAGGAGGGGGCCGAGC
 GCCCCGAGAAACGCCCTCGCCCTGTGAGGGCTGAATGGGGCTGAGATAGAGGGCTCTGCCCTGTCTGGG
 GCCCCCGGAAACCCCTCAGAGATGGCCACAACTCACTGCCCAATGGCCCGGCCCTTGACAGCGGC
 CAGCCCCGGTGTCCCCGCTGAGCCAAGCCCTGAGAAAGTGGCCACCATCTCCCCAGACGCCATGTCTAT
 GCCAGGCTCTCGCCCCAGGAGCCGTATCCCTGTCTCTCTCTGAGGAGGACACGGGCTCGGAGCCCTCA
 GGCTCACTGTGCGCCAAAGAGCGGTGGAGCAAGCGGGCTCGGCCGAGCAGGACCTGGCGCGGCTGGTGA
 TGGAGAAGAGGCAGGGCGCCTGTGTGCGGCTGGCCTCAGGGGCTCGTCTCTCTCCAGTGAGGAGCA
 GCGCCGTGCTCTGAGACCTCTCAGGCACGGGCTCTGAGGAGGACACGCCCGCCTCTGAGCCGGCAGCG
 GCCTTGCCCAGGAAGAGCGGGGAGGGCAGCCGCCACCAGGAGCCGGATTCCCCGCCCATTTGGCCTCCGCA
 TGCCCATGCTGTGTCAGCCAGCAGCCCGCAGCAGATCCCATGGCGCGGCCCCAGCATTTGGACACAGC
 CATCACCAGCAGGCTCCAGCTGCAGACGCCCCAGGGTGGGCCACTGCTGCTGACCTCCGCCCCAAACAA
 CCTCTTGGCCGCGGCTGGGCCAGGGCGAGCCCAAGCCGGAGCCAGGCCCCAGCGCCGCGCAGCCCGC
 GCCTCCCCGCGTCCACATCCGCCGCGCGCAATGCCAGCGCTCCCCCGGAGCCAGTCCCTGTCCCGCAG
 AGAGAGCCCCCTCCCCCTCGCACCAGGCCCGGGCGGGTCCCCCGCCCCGGGGCGTCCCGCCGGCCCGG
 GCCCAGCTGATGGCACCCCCCTCCCCGGGGGCTCCAAGAAAGGACCCAGAGGGAACCTCCAGGCTCAGC
 GCCCAACCAAGGAGGCGGAGGCGCGGAGGGCCGGGCTGGGGCCAGATAATGACGCCGCTGTCT
 CTCCGCGGTCCCCCACCCTCACCCTCGCCCCCACCCTCGAGCCGGCCACACTGGAGCAGCTCCAGCACA
 GCCTTACGCGCCGACGCGCGCCACCCGCGGCCCCAGCTTTCCGCTGACCCGCGAGGACGCGCGCGAG
 CACACGCGGCGCCCCGCCAGGCCTTAGGGCCGTGGGGGACGCGGCCCGCGCCGCGGGGAGGGTCTGCC
 TCCCCCTCTCGCCCTGTGTCTCTCATCTCCCGCCGCCCGTCAGGCCGGCCAGCCTCACATCAGTCTC
 TCCGCCCGGGGAAGGCTCAGCCACTTTTCACTGAGGACTCCACTTCTGGGGACGCTGGTTCTGCTGCC
 CACCAGGCCTAGGCTACGCTCCATGCTCCCCAGCAATCTCTGCCTACACCTCTGCGCGCCTTGCCCT
 CCTCCGACCCCTTTCCAGCCAAAGTCCCCCACCCTTCCAGAGAAGCAGCCTCAAAATCCAGAAGTGGAG
 GCTCCAGCCTCCCCGCGAGGGTCCAGCCCCACAGTCTTCTGGGAGCCATTGTGGCCAGGGACGGCCTCTG
 GACTGCCAGGCTGGGTGGGGACCCAGGGAACATCGGTCTACTCAGGTGTGAGGGGGCAGGTCTGACCTG
 CCCAAAGTTGGCTCCATCTTGGAACCTCGGTGAGAGGCAGTGGGCAAGTGATCTTGAGATGGGTGGG
 CAGGTGATCTGTGGGCAGGGGATGTGCTCCCTGACACTCTGGGGTGCAGAAACCTCTTGCTCCAGAT
 TTGGGTGGAGCCTCTGTGGGAACCATAGGAAGTGTGTGGGCTGCCTTCTTGGGCAAGTATTTCCAGTG
 GAAGTTGGAGGGGGCTTTAAACAAAGTTTACTCCCTCCCTGTCTCCCTGATCTAGTGCTCAGGACCCTT
 CACCATCAGGAATTCCTTCTGTCTATCTAACCTCAGTCTGCTTACTGCAGTTCCAGCCAACTGCTCTT
 TCTGAGTTCAAAGCAGGTGGAGACTGGCTGGTTACCATCTTTGCACTGGCCCTTCGGAGATTCGGGGAC
 TCAGTTCTGGTGGGGTCACTCTCCCTGTCTCCCGCTGTGGGAGGGAGGGAGGGCTGGCTCAGGCATCG
 TCTCCCGCAATGGGCAGAGAGAGCAGAGACAGGTGGACCAACAGACAGCTGGCCCCCTGGAGGCAGAAAG
 CCCTTCTAACTTCCAGATTGTATGCTTGAGTGATGGGTCCCCAGCCCAAGCCACTCTTCCCTCAGCTCA
 CCCTTCCAGCTGTCTCTTCTTGCCCTGACCCAGCCCGTGCAGCTGCTCTACTCCAGGAATGGATGTGGG
 GACTCTTCTGGGTCTTGGCTCCTGCATAGCTCACCACCTCATCATGAGCCTCACTGCCTACATCTG
 GGGCAAGCAGCACACCGGCTGCAGATGGGACAGCCAGCCCTGCTATCTGGACAGGCCCTTGACGCTCT
 GTCCCCCTGAGCTCTGTCTCTTCTTCTGAGTACAGAGAGCAAGCAAGACATCCAGGGAAGAGG
 AAGAAAGGCCCTTAGTCTCCCGCAGCAGTCTGCTGCGTCCAGCCACCATCACCCTGGGAAGGATGCCACAA
 GGCAGCTGACCTGAAAGCAGCCTCCCCCTCATGGAGAGTCAGCAGCTTGGGCAGCCACTTCCAGGCCAG
 GGTGGTGGCTTCTCTGCAGACCAGCTGAGGGGAGGACTCCTGGGTGGACAGCCTTTGACGTCCACCCAC
 GCTGATGCAGAAGCTCCAGAACACTCAGGAACTTCTCCGACAGAGCCCTCCTTGCTCACTTGAGGCC

CTCCCAAGGCCCTCTACTGCCCTCTGGGTCCAGCAGAGGGAGTGGAGGAAGGGCCACTGCCTCCCACCTA
 GAGCTTCTCCGAATGACAATCAGCTCGTGCCAGGTGGGGACCAGGATATGACTCCTGGTGCCAGGCCCT
 GGGCCTGCTCCTTGCCACCAACCGAACCCTGAATGTAGGGCCCCAGCCTCACCTCTGCCCCAGGACCAA
 CAACACCCCTGGTTTGGAGCTGGGAGGAAGAAGGGGGCTTGAGAGAGCCCCAGGTCCATTCTACCCCCAGC
 TTCCTCAGCACTGGAGCTGGCAGAGACGAAAACCCAGTCTGCCCTTGGGATTCCAAACCTCCCTAGGG
 CTCCCAACTGACCTCAGGCCCTCTGAGTCACTGAATGTCACCAGGAGAGGTGGGGGAGGAAAGTGGGCCA
 CTGGGGAGGGGTACCTAGGGGACTGCCTCTGTGCCCTCTCCCCAGGAAGCATCCAGGGCAGAGGAAGCC
 ACATCTCCCGGTGCCCCCAACCCAGCTGCAGCCTCCTCCCCCTGAGCATTCAATCTCTCCACCAGGCCCT
 CCAGGTCTTGAGCCCTTCTCTGTAAAAGTGTACACCACCTCCCTCAGCACTTCCCCATCACAACAACC
 TATGTCACTGACTCAGATGCAGGGTCTGCTCACCCCAACACATGCCTTCCCTCCCCAGCCACACCGTGCA
 CGAAGGGGGACAGGAGAGGAGAGGGGCTGTGCCCCAGGCTCCCCATTTCCAGCTCTTCACAGAGGCCCT
 GGTTTGTCTAGTCTTCTGAACTCCAGGACCAGCCCTGGTGGCATGGGGTGGGGAGCAGGGAGTTGCC
 TTCCCTTCCCTCGGAAGCCACCTAAGAATGTTTACATGCCAAACAGAATGTAACACCCCTCCCCAAGCC
 CTTCCAGTCACTGCATGGCCTCTGCCCATCCTGCACCTGTCCACCCCAACCCCTGGAAGCCAC
 TGTCAATGATTAGATCGGGTCTCGGAAGGGAAGTAGCCATCACACCATTAAGGCTGTGGACCTTT

>gi|28466990|ref|NM_173500.2| Homo sapiens tau tubulin kinase 2
 (TTBK2), mRNA

AGCAGGTGCTGGCACAAGAGCAGCGGCTTGGGGGAGCCGGCAGCAGCAGTAACAGCAGCAGCAGCCGCCG
 CCGCCGCCGCGAGTAAACCGCGGACGGTACCCAGGGGACTACCCAGCCGGCCGGCCCTGGAAGCCGCGCT
 CGGGTCCCGCCGAGCTCGGCGGTGGGGGATGGGCAGGCAGTGGCGGTCCCGCTGCCGAGGGTTAACCCC
 CGCCGGTCCCGGTCTGAGCTGGACCAGAGCCCTCCTCCAGAAACCCCTGCGTCCGCCACGGCCAGGTT
 AAATGGAAACCACCCCTTGGGAACGGATGCCTGTGTAGCTGTCTACCATATCAGTGTATTGCAATGAGT
 GGGGGAGGAGAGCAGCTGGATATCCTGAGTGTGGAATCCTAGTGAAAGAAAGATGGAAAGTGTGAGAA
 AGATTGGGGGTGGGGGCTTTGGAGAAATTTACGATGCCTTGGACATGCTCACCAGGGAAATGTTGCACT
 GAAGGTGGAATCAGCTCAACAACCAAAACAAGTTCTGAAATGGAAGTTGCTGTTTTGAAAAAGCTGCAA
 GGAAAGACCATGTTTGTAGATTTATTGGCTGTGGGAGGAATGATCGATTCAACTATGTGGTCATGCAGTT
 GCAGGGTCGGAATCTGGCAGATCTTCGCCGTAGCCAGTCCCGAGGCACATTCACCATAGTACCCTCTC
 CGGCTGGGTAGACAGATTTTGGAGTCTATTGAAAGCATTCAATCTGTGGGATTCTTGCATCGAGACATCA
 AACCGTCGAATTTTCGATATGGGTGCTTTCTCTAGTACATGTAGGAAATGTTACATGCTTGATTGGCTT
 GGCTCGACAATTTACCAATTCCTGTGGTGACGTACAGCACCTCGAGCTGTGGCAGGTTTTCGAGGACA
 GTTCGTTATGCATCAATCAACGCACATCGGAACAGGGAAATGGGAAGACATGATGACCTTTGGTCTTTAT
 TCTACATGTTGGTGGAGTTTGTGGTTGGTCACTGCCCTGGAGAAAAATAAAGGACAAGGAGCAAGTAGG
 CTCTATTAAGGAGAGATATGACCACAGGCTCATGTTGAAACATCTCCCTCCAGAATTCAGCATCTTTCTA
 GACCATATCTCTTCTTTGGATTATTTTACAAAACAGACTACCAGATGTCCATCAGGGTGACCCGGAAGT
 CCTACAAGGTGTCCACCTCTGGCCCCCAGGCCTTTAACAGCCTCTCCTACACGAGTGGGCCCTGGTGCCCTG
 CATCAGCTCCTCGAGCTTCTCCGAATGGGCAGCAGCAGCTTCCGGGTGGCCTGGGTGCAGGATATGGT
 GGGGCCAGTGGAGGCATCACCAACATCACTGTCAACCAAGCCTGCTGAGCCCTCTTAACCTGGAGGTGG
 ACCCCAACATCCAGGCCGTGCGCACCCAGGAGGAGAAGCAGATCAAGACCCCTCAACAACAAGTTTTTCTC
 CTTCATAGACAAGGTACGGTTCTTGGAGCAGCAGAAACAAGATGCTTGAGACCAAGTGGAGCCTCGTGCGAG
 CAGCAGAAGATGGCTCGGAGCAACATGGACAACATGTTTCGAGAGCTACATCAACAACCTTAAGTGGCAGC
 TGGAGACTCTGGGCCAGGAGAAGCTGAAGCTGGAGCGGAGCTTGGCAACATGCATGGGCCTGGTGGAGGA
 CTTTCAGGAACAAGTATGAGGTTGAGATCAGTAAATGTACAGAGATGGAGAATGAATTTGTGCTCATCAAG
 GAGTATGTAGATGAAGCTTACATGAACAAGATGGAGCTGGAGTCTTCCCTGAAAGAGCTGACTGCCAAGA
 TCAGCTTCTCAGGCAGCTGTATGAAGAGGAGATCGGGAGCTGCAGTCCCAGATCTCGGATACATCTGTG
 GTGCTGTTTCATGGACAACAGCCGCTTCTTGGACATGGACAGCATCATCCCTGAGGTCAAGGCCAGAGGGC
 TTCCCTGGAGGCTGCCACGCAGATACCGAGCAGCGTGGGGAGCTAGCCATTAAGGATGCCAATGCCAAGC
 TGTCTGAGCTGGAGGCCGCCCTGCAGCTAGCCAGTCAAGACATGGCGCGGCAGCTGCGTGAGTACCAGGA
 GCTGATGAACGTCAAGCTGGCCCTGGATATCAAGATCGCCACCTACAGGAAGCTGTTGGAGGGCGAGGAG
 AGCTGGCTGGAGTCTGGGATGCAGAACATGAGTATCCATATGAAGACCACCAGCAGCTATGCAGGTGGTC
 AGAGCTTGGCCTATGGGGGCTCACAAGCCCTGGCCTCAGCTACGGCCTGGGCTCCAGCTTTGGCTCTGG
 CATGGGCTCCAGCTCCTTCAGCCACACAGCTCCTCCAGGGCCGTGGTCATGAAGAAGATCGAAACCCGT
 GATGGGAAGCTAGTGTCTGAGTCTTCAACGTCCTGCCAATTTGAGAATTACTGCCACACTTCTTACAT
 CCGTGTGTTGACAAATGACATCAAGACTTTTGGAGTAATTTGAGAGTGACCCCTTTTGACTGGGAGAAAGACTGG
 AAATGATGGCTCCCTTAACAACCAACCACTACTTCTACCACCCCTCAGTTGCACACTCGCTTGACCCCTGCT
 GCAATTGGAATTGCCAATGTACTCCCATCCCTGGAGACTTGTCTCGAGAAAATACAGATGAGGTATTTTC
 CAGATGAACAGCTTAGCGATGGAGAAAATGGCATCCCTGTTGGTGTGTACCAGATAAATTGCCTGGATC

TCTGGGACACCCCCGTCCCCAGGAGAAGGATGTTTGGGAAGAGATGGATGCCAACAAAAACAAGATAAAG
 CTTGGAAATTTGTAAGGCTGCTACTGAAGAGGAGAACAGCCATGGCCAGGCAAATGGTCTTCTCAATGCTC
 CAAGCCTTGGGTCAACCAATTCGTGTCCGCTCAGAGATTACTCAGCCAGACAGAGATATTCCTACTGGTGCG
 AAAGTTACGTTCCATTACAGCTTTGAGCTGGAAAAACGTCTGACCCCTGGAGCCAAAGCCAGACACTGAC
 AAGTTCCCTTGAGACCTGCCGAGAGAAAATGCAGAAAGATACCAGTGCAGGAAAAGAATCTATTCTCCCTG
 CTCTGCTGCATAAGCCTTGCGTTCTGTGTGTCCCGTACTGACCACATCTGGCACTATGATGAAGAATA
 TCTTCCAGATGCCTCCAAGCCTGCTTCTGCCAACACCCCTGAGCAGGCAGATGGTGGTGGCAGCAATGGA
 TTTATAGCTGTAACTGAGCTCTTGCAAGCAAGAAATGATTCCAAAGAATGGGTGATTGTGGACAAGG
 AGCAGGACCTTCAGGATTTTAGGACAAATGAGGCTGTAGGACATAAAACAACCTGGAAGTCCTTCTGATGA
 GGAGCCTGAAGTACTTCAAGTCTTGAGGCATCACCTCAAGATGAAAAGCTCCAGTTAGGTCTTGGGCA
 GAAATGATCATTTAAAGAAGGAAACCTCAGGTGTGGTCTTAGCACTTCTGTCAGAGGGTCTCTCTACTG
 CTGCTTCAGAACAAATATACAGATAGGCTGGAACCTCAGCCTGGAGCTGCTAGTCAGTTTATTGCAGCGAC
 GCCCACAAGTCTAATGGAGGCGCAGGCAGAAGGACCCCTTACAGCGATTACAAATTCCTAGACCTTCTGTG
 GCATCTACACAGTCAACTTCAGGAAGCTTTCACTGTGGTGCAGCAGCCAGAGAAGAAAGATCTTCAGCCCA
 TGGAGCCCACTGTGGAACCTTACTCTCCAAGGGAACCTTCTCTGGCTTGGTTGTGACAGAGGGTGAACC
 TCCTAGTGGAGGAAGCAGAACAGATTTGGGGCTTCAGATAGATCACATTGGTCATGACATGTTACCCAAC
 ATTAGAGAAAGTAACAAATCTCAAGACCTGGGACCAAAAGAACTTCTGATCATAATAGACTGGTTGTGA
 GAGAATTTGAAAATCTCCCTGGGGAACCTGAAGAGAAAAGCATCCTTTTAGAGTCAGATAATGAAGATGA
 GAAGTTAAGTAGAGGGCAGCATTGTATTGAGATCTCCTCTCTCCAGGAGATTTGGTAATTGTGGAAAAG
 GATCACTCAGCTACTACTGAACCTCTTGATGTGACAAAAACACAGACTTTTAGTGTGGTGCCAAATCAAG
 ACAAAAATAATGAGATAATGAAGCTTCTGACAGTTGGAACCTTCAGAAATTTCTTCCAGAGACATTGACCC
 ACATGTTGAAGGTCAGATAGGCCAAGTGGCAGAAATGCAAAAAATAAGATATCTAAGGATGATGACATC
 ATGAGTGAAGACTTGCCAGGTCTCAAGGAGACCTCTCTACTTTTTTGACCAAGAGGGCAAGAGAGAGA
 AAATCACCCCTAGAAATGGAGAATATTTCAATTGTGTTTCAGAGAATGAACATGGTGCCCCAACCCGGAA
 GGATATGGTTAGGTCATCCTTTGTAAC TAGACACAGCCGAATCCCTGTTTTAGCACAGAGATAGACTCA
 ACTTTGGAATCATCCTCTCCAGTTTCTGCAAAAAGAAAAGCTCCTCCAAAAGAAAGCCTATCAGCCAGACC
 TAGTCAAGCTTCTGGTGGAAAAAGACAATTCAAGTCCTTCTTGGCGACCTCTCAAGTGCCCTCTGATAA
 ATTGCTAGAGGAGAAACTAGCTACTGTTCCTGCTCCCTTTTGTGAGGAGGAAGTGCTCACTCCCTTTTCA
 AGACTGACAGTAGATTCTCACCTGAGTAGGTGAGTGAAGATAGCTTTCTGTCAACCATCATCTCCAGT
 CTAGAAAGAGCAAAATTCAGGCCAGTTTCATGGGTCAACACAGATCAGGTCAATAGCTCAACTTCGTC
 TCAGTTCTTTCTCGGCCACCACCAGGAAAGCCACCCAGGCTTGAGTAGAAGCCAGGCTACGAGAGA
 TATAAAGTCCTAGGGAGTAGTAACCTCCGACTCAGACCTTTTCTCCCGCCTGGCCCAATTTCTTCAAAATG
 GATCTCAGAAACCCCGGAGCACTACTCAGTGCAAGAGTCCAGGATCTCCTCACAATCCAAAAACACCACC
 CAAGAGTCCAGTTGTTCCTCGCAGGAGTCCCAGTGCTCTCTCGAAGCTCATCTTGCCCTCGCACGTCT
 AGTTCTCTACCATCTAGGGCTGGACGGCCCCACCATGACCAGAGGAGTTCTGTCACCCATCATCTCCAGT
 GCAAGTCACTCCAGCCACTCAGGATCTTCTCTCCAGGAGGTCTGCCAACAGGAGCATTGCAAACC
 CAGCAAGAATGGCTGAAAGGATCCGGCAGCCTCCACCACCACTCAGCCAGCACTAAACCCCCCAAGGG
 AAGAGTAAGCCAGCCAGTAACTCAGCAGATAGGAGCCAGGCTGCATCTCTTTGAAAGGTGTGAGATCTT
 CCTCTAAACCTGATGCATGTGTGTCCTGTACTTTCTATGTAAAAAATCAGTGTGATCTTCTCTTGC
 AAAAGAAAGTAACATGATCAATTATTTATAAGAAGACATAATACATGATAAGGAATTACCTAAGGCAGGC
 AGCAAGTAGATTAGGAATCAATGCTTTGTACAAGAAGGAAAAATAGAGCAAAAATCCAAGGGGGAGAAA
 CTCATTAAATGA

>gi|4505466|ref|NM_002526.1| Homo sapiens 5'-nucleotidase, ecto (CD73) (NT5E), mRNA

GCACTCGCCCGGCTCGCCCGCTTTTCGCACCCAGTTCACGCGCCACAGCTATGTGTCCCCGAGCCGCGCGG
 GCGCCCGGACGCTACTCCTCGCCCTGGGCGCGGTGCTGTGGCTGCGGCTGGCGCCTGGGAGCTTACGA
 TTTTGCACACCAACGACGTGCACAGCCGGCTGGAGCAGACCAGCGAGGACTCCAGCAAGTCCGTCAACGC
 CAGCCGCTGCATGGGTGGCGTGGCTCGGCTCTTACCAAGGTTTCAGCAGATCCGCCCGCGCCGAACCCAAC
 GTGCTGCTGCTGGACGCGCGGCGACCACTACCAGGGCATTATCTGGTTTACCCTGTACAAGGGCGCCGAGG
 TGGCGCACTTCATGAACGCCCTGCGCTACGATGCCATGGCACTGGGAAATCATGAATTTGATAATGGTGT
 GGAAGGACTGATCGAGCCACTCCTCAAAGAGGCCAAATTTCCAATTTCTGAGTGCAACATTTAAAGCAAAG
 GGGCACTAGCATCTCAAATATCAGGACTTTATTTGCCATATAAAGTTCTTCTGTTGGTGATGAAGTTG
 TGGGAATCGTTGGATACACTTCCAAAGAAACCCCTTTTCTCTCAAATCCAGGGACAAATTTAGTGTTTGA
 AGATGAAATCACTGCATTACAACCTGAAGTAGATAAGTTAAAAACTCTAAATGTGAACAAAATTTATTGCA
 CTGGGACATTCCGGTTTTGAATGGATAAACTCATCGCTCAGAAAGTGAGGGGTGTGGACGTCGTGGTGG
 GAGGACATCCAACACATTTCTTTACACAGGCAATCCACCTTCCAAAGAGGTGCCCTGCTGGGAAGTACCC

ATTCATAGTCACTTCTGATGATGGGCGGAAGGTTCCCTGTAGTCCAGGCCATATGCTTTTGGCAAATACCTA
 GGCTATCTGAAGATCGAGTTTGATGAAAGAGGAAACGTCATCTCTTCCCATGGAAATCCCATTTCTCTAA
 ACAGCAGCATTCCTGAAGATCCAAGCATAAAAGCAGACATTAACAAATGGAGGATAAAAATTGGATAATTA
 TTCTACCCAGGAATTAGGGAAAACAAATTGTCTATCTGGATGGCTCCTCTCAATCATGCCGCTTTAGAGAA
 TGCAACATGGGCAACCTGATTTGTGATGCAATGATTAACAACAACCTGAGACACACGGATGAAATGTTCT
 GGAACCACGTATCCATGTGCATTTTAAATGGAGGTGGTATCCGGTCGCCCATTTGATGAACGCAACAATGG
 CACAATTACCTGGGAGAACCCTGGCTGCTGTATTGCCCTTTGGAGGCACATTTGACCTAGTCCAGTTAAAA
 GGTTCACCCCTGAAGAAGGCCCTTTGAGCATAGCGTGACCCGCTACGGCCAGTCCACTGGAGAGTTCCCTGC
 AGGTGGGCGGAATCCATGTGGTGTATGATCTTTCCGAAAACCTGGAGACAGAGTAGTCAAATTAGATGT
 TCTTTGCACCAGTGTGAGTGCCAGTTATGACCCCTCTCAAAATGGACGAGGTATATAAGGTGATCCTC
 CCAAACCTTCTGGCCAATGGTGGAGATGGGTTCCAGATGATAAAGATGAATTATTAAGACATGACTCTG
 GTGACCAAGATATCAACGTGGTTTCTACATATATCTCCAAATGAAAGTAATTTATCCAGCAGTTGAAGG
 TCGGATCAAGTTTCCACAGGAAGTCACTGCCATGGAAGCTTTCTTTAATATTTCTTTCACTTTGGGCA
 GTGATCTTTGTTTTATACCAATAGCCAAAATTCCTCTTGCCCTTTAATGTGTGAAACTGCATTTTTTCAA
 GTGAGATTCAAATCTGCCTTTTAGGACCTGGCTTTGTGACAGCAAAAACCATCTTTACAGGCTCCTAGAA
 GCTGAAGGTTAGAGCATTATAAAATGAAGAGACAGACATGATTACTCAGGGTCAGCAACCTAGTGAGTTA
 GAAAAAAATTAACATAGGGCCCTATAAGGAGAAAGCCAACATATGTTAAGTTTACGTGTCCAAATTTTAA
 TGAAATTTTACTAACAATTTTAAACCATATTTTTCTTCTCATATCCATTTCTAATCCATCAAACAGCTT
 ATGTTTACATAAAATTTTATCATTCAAGGAAGTTTTAAGCACACTGTCTCATTTGATATCCCAACTT
 ATTTTGGTAGGAAAGAGAGATGTTTTTCCACCTGTGATGAAAAAACTGAAGCTCAAAAAGGTTGA
 TTTGACCATACAGCTAATGTGACAGATCCAAGACCTAGACCTAGGTCTTTTGAAGTCAAGTCCAGCATT
 CTCAACTATATCAAGTTACTGTTTCAAGTAATCTCTCTCTTCTATAATTATCAATAGCCCCAAG
 CTCATGGATGACAAATCTCTGCTTTATTTCTTGTCTCTATTTTTTCACTTTATAGCTCCTGTTATAATAG
 CAAGTTTAAATGGTATAAACACAGGATACCATCCTCTCTTGCAACACCCATGTGCCCTTGATGAGTCAGGT
 AGCAAGCTGTAGTAGATAATGAGAAAGGCCAGAGGCTGCAAGACAGTCAAAGGACACGAGAGAAAGGAA
 GGGGAAGAACAGGACTCCAGGACTGTTTTATATTATAGAAAAGCAAGAGCTAAAGAGCATTTACACATGT
 TAAACAGATACTTGTTAAGCATAGTGCTGACACACGGCATTAGCTGTTATTTTATGAGATTCCATCAGC
 TCTGCCTCTGTCTCTTTCTTCTAACATGAAGGTATCATGAGAAGAGAACCTTCTAACATAAGCTGTAAT
 TCTAAACCTGCCTTGTCCCTCTCCAGCAAGAGGCTAGCACTGAATTCATTCTACTCATACTACACACCC
 AGTTATGGAATGTCCAGAGTTCTCGAAGAAAATAAATGACTTTAGGAAGAGGTATACATTTTTTAAAGTCG
 CTCTGCCTCCAAATCTGAACAGTCACTGTAATCATTCTTAAAGCCAGATATGAGAACTTCTGCTGGAAA
 GTGGGACCCCTGAGTGGGTGGTCAGAAAATACCCATGCTGATGAAATGACCTATGCCCAAAGAACAAT
 ACTTAACGTGGGAGTGGAACACATGAGCCTGCTCAGCTCTGCATAAGTAATTCAAGAAATGGGAGGCTT
 CACCTTAAAAACAGTGTGCAATGGCAGCTAGAGGTTTTGATAGGAAGTATGTTTGTCTTAGTGTTTA
 CAAATATTAAGTACTCTTGATACAAAATATACTTTTAAACTTCATAACCTTTTTATAAAAAGTTGTTGCAG
 CAAAATAATAGCCTCGGTTCTATGCATATATGGATTGCTATAAAAAATGTCAATAAGATTGTACAAGGAA
 AATTAGAGAAAGTCACATTTAGGGTTTATTTTTTACACTTGGCCAGTAAAATAGGGTAAATCCTATTAGA
 AATTTTTTAAAGAACTTTTTTTAAGTTTCTTAAATCTGTGTGTGATTGTGAAGTGGTATAAGAAATGAC
 TTTGAACCACTTTGCAATTGTAGATTCCTCAACAATAAAATTTGAAGAT

>gi|17402894|ref|NM_002876.2| Homo sapiens RAD51 homolog C (S.
 cerevisiae) (RAD51C), transcript variant 2, mRNA
 GTGCGGAGTTTGGCTGCTCCGGGTTAGCAGGTGAGCCTGCGATGCGCGGGAAGACGTTCCGCTTTGAAA
 TGCAGCGGGATTTGGTGAGTTTCCCGCTGTCTCCAGCGGTGCGGGTGAAGCTGGTGTCTGCGGGGTTCCA
 GACTGCTGAGGAATCCTAGAGGTGAAACCTCCGAGCTTAGCAAGAAGTTGGGATATCTAAAGCAGAA
 GCCTTAGAACTCTGCAAAATTATCAGAAGAGAATGTCTCACAATAAACCAAGATATGCTGGTACATCTG
 AGTCACACAAGAAGGTACAGCACTGGAACCTTCTTGAGCAGGAGCATACCCAGGGCTTCATAATCACCTT
 CTGTTCAAGCACTAGATGATATTCTTGGGGGTGGAGTGCCCTTAATGAAAACAACAGAAATTTGTGGTGCA
 CCAGGTGTTGGAAAAACACAATTATGGTAAATAAAGTGTCTCCTTTTAAAGGGTGGGTTAATAACATA
 TTATGAAAGTAGTATTTGTACTATCGTCAGGAAACCAATAAGATATATATGTGCTCTTAATTTTAAAGT
 GTGTATGTGCATTAACAAAAAATTAGCTTACAAAAA

>gi|4885656|ref|NM_005431.1| Homo sapiens X-ray repair complementing
 defective repair in Chinese hamster cells 2 (XRCC2), mRNA
 TTTGACTGGCCGTAGAGTCTGCGCAGTTGGTGAATGGCGTTGGTGGCGGGAAGTTGAGTCTCTCCTGCG
 CCGAGCCTTCGGGGCGATGTGTAGTGCCCTCCATAGGGCTGAGTCTGGGACCGAGCTCCTTGCCCGACTT
 GAAGGTAGAAGTTCTTGAAGAAATAGAACCAAACTGTTTGTGCTGATGAAGATTCACCTGTGCATGGTG

ATATTCTTGAATTTTCATGGCCCAGAAGGAACAGGAAAAACAGAAATGCTTTATCACCTAACAGCACGATG
 TATACTTCCCAAATCAGAAGGTGGCTTGAAGTAGAAGCTTATTTATTGATACAGATTACCACTTTGAT
 ATGCTCCGGCTAGTTACAATTTTGAGCACAGACTATCCCAAAGCTCTGAAGAAATAATCAAATACTGCC
 TGGGAAGATTTTTTTTGGTGTACTGCGAGTAGTAGCACCCTTACTTCTTACACTTTACTACTAGAAAG
 TATGTTTTGTAGTCACCCATCTCTCGCTTTTGATTTTGGATAGCCTGTGAGCTTTTACTGGATAGAC
 CGCGTCAATGGAGGAGAAAGTGTGAACCTACAGGAGTCTACTCTGAGGAAATGTTCTCAGTGCTTAGAGA
 AGCTTGTAATGACTATCGCCTGGTTCTTTTGAACGACACAAACTATAATGCAGAAAGCCTCGAGCTC
 ATCAGAAGAACCTTCTCATGCCCTCGACGACTGTGTGATGTGGACATAGACTACAGACCTTATCTCTGT
 AAGGCATGGCAGCAACTGGTGAAGCACAGGATGTTTTTCTCCAAACAAGATGATCTCAAAGCAGCAACC
 AATTTTCATTAGTTTCACGTTGTTTTAAAAAGTAACAGTTTAAAAAACATTTTTTTATTATTGGAGAAAG
 TGGGGTTGAATTTTGTGATATACATCAAAAATAGTCTTTTGCAGGGTACTACGCAAGCCTTAAATTTT
 TTCTTAAGACAGAGTCTTGCTCTGTCTCCAGGCTGGAGTGCAGTGGCACAATCATGGCTCACTGCAGCC
 TTGAACCTCTGGCTCAAGGGATCCTCTATGTGTGCCCTCTAGAGTGCAGGGATTACAGGCGTGAGCCA
 CTGCTCGTGGCCAAAAGTTTTCTTTTTTTTTTTTTTTCTTTTGAAACAGTCTTACTCTGTCTCCAGGC
 TGCTGGAGTGCAGTGGCACAATCTCGGCCCGCTGCAGCCTCTGCCTCTTGGGTTCAAGTGATTCTTCCAC
 CTCAGCCTCCCAGGTAGCTGGGATTACAGGCACCCACCACCACGCTGGCTAATTTTTGTATTTTTAATA
 GAGACGGGGTTTCAACCATGTTGGCCAGGCTGGTCTCGAATCTTGACCTCAAGTGATCCACCCACCTCGG
 CCTCCCAAAGTGCTAGGATTACAGGCCCGTGCCAGCCCTAAAGTTTAACTCTAGGGGAATTAACAGT
 ATTTCTTTACAGAATGGATTTGTAACTAGCACAGTAAAGTAAAGACTATTCTGTTTCTAGGCTGTG
 AATCAAAGTGATTTAGCAATTAACCTTTGTATTAATTTACCACCAATATTTCTTCACAAAGGAACCTTTT
 AAAAGATTATCTCAGAAAGTAAATCTGAGAGGTAAGAAGTAATAATGAGTAAATGGTAAGTACTTGAGTA
 AATCTAAAGAAATATTGATAGTAAGGCAATCCTAAGCAAAAAGAAACAAAGCTGGAGGCATCACGCTACCC
 AGCTTCAAACATACTACAAGGCTACAGTAACCAAAACAGCATAGTACTGGCACAACAAACACAGTAGAC
 TGATGGAACAGAATAGAGAATTTAGAAATGAGACCACACCTATAATTTTTTTGATCTTCGATGAACCT
 GACAAAAACAAGCAATGGGCAATGGATTCTCTATTCAATAAATCGTGCTGGGATAACTGGCCAGCCATAT
 GGAAAAGATTGAAAATGGACGCCCTTCTTATGCCATATACAAAATTAACCTCAAGATGGATTAAAGACTT
 AATGTAAACCCCAAAACAGTAAAAATCCTGGAAGACAACCCAGGCAGTACCATTACAGGACATAGGCACAG
 GCAAAGATTTTCATGACGAAGACGCCAAAAACAATTGCAACAGAAGCAAAAATTCACAAATGGGATCTAAT
 TAACTAAAGAGCTGCACAGCAAAAGAACTATCAAGAGAGTAAACAGACAGCTTACAGAATGGGAGAAA
 ATTGTTGCAACTATGCATCTGAGAAAGGCTGAAATCCAGCATCTATACGTAATTTAAACAAATTTAGA
 AGAAAAAACCCCATTAAGAAAGTGGGCAAGGACATGAACAGACACTTTTCAAAGAGACATCTGTG
 GCCAAGATTCCTATGGAAAAAAGCCAGCATCACTGATCAATTAGAGAAATGCAAAATCGAAACAACAACGA
 GATACCATCTCACACAGTCCAAATGGCTATTATAAAAATGTGAGAAAATAACAGATGCTGGTGAGGTTG
 TGGAGAAAAAGATATGCTTATACACTGTTGGTGAAATGTAAATTAATTAGTTACAGCCATTGTGGAAGA
 CAGTGTGGGGATAAAGACAGAGATACCATTCAACCCAGCAATCTCATTACTGGGTATATACCCAAAGGAA
 TAGAAATCATTTGTATAAAGACACATGCACGCGTATGTTCTGTCAGCACTGCCCATCAGTGACAGACTG
 GATTAAAAAATGTGGTACATACACACCAGGAATACTATACAGCCATAAAAAGGAACAAGACTGACTGG
 GCGTGGTGGCTCATGCCTGTGATCCTAGCACTTTGCGAGGCCGAGGTGGGTGGATTGCCCGCGCTCAGGA
 GGTCAAGACCAGCCTGGGCAACACGGTGAACCCCATCTCTATTAAAAATACAAAAATTAGCTGGGCATG
 GTGGTGCCTGCTGTAGTGCCAGCTACTCAGGAGGCCGAGGCAGGAGAATTGCTGGAACCCAGGAGGTGG
 AGGTTGCAGTGAGCTGAGATCGCGCCATTGCACTCCCGCTGGGCGACTCCATCTCTAAAAAATAAAAAA
 AAAAAA

>gi|11968151|ref|NM_016148.1| Homo sapiens SH3 and multiple ankyrin repeat domains 1 (SHANK1), mRNA

GTCGCCCCGTGGCCCCACAATGACCCACAGCCCCGCGACAAGCGAGGACGAGGAACGCCACAGTGCCAGC
 GAGTGTCCCAGGGGGGCTCAGAGTCCGACAGCTCCCCAGACGGGCCAGGTGAGGCCCCCGGGGACCC
 GGGGCCAGGGCAGTGGGGCACCCTGGTAGCCTGGCCTCTGTTAGAGGCCTCCAGGGCCGCTCAATGTCCGT
 CCCAGACGACGCCACTTCAGCATGATGGTCTTCAGGATTGGCATCCCGGACCTGCACCAGACAAAATGC
 CTTCGCTTCAACCCCGATGCCACCATCTGGACGGCCAAGCAGCAGGTGCTCTGTGCCCTGAGCGAGAGCC
 TGCAGGATGTGCTCAACTATGGCCTGTTCCAACCGCCACCTCCGGCCGCGATGCCAACTTCTGGAGGA
 GGAGAGGCTGCTGCGGGAGTACCCCACTCTTTGAGAAGGGGTCCCCTACCTGGAGTTCCGATACAAG
 ACCCGAGTTTACAAACAGACCAACCTGGATGAGAAGCAGCTGGCCAAGTTGCACACGAAGACGGGGTTGA
 AGAAGTTCTTGAGTATGTGTCAGCTCGGGACATCTGACAAGGTGGCGCGGCTGCTGGACAAGGGCTTGA
 CCAATTAACCATGACTCGGATTTCGGGAGAGACCCCTTGACACTGGCGGCCAGACCGAAGGCTCTGTA
 GAGGTGATTGCAACCTGTGCCTGGGCGGGGCCACATTGACTTCCGGGCCCGGATGGCATGACCGCAC
 TGCATAAGGCCGCATGCGCCGACACTGCCCTGGCACTCACGGCGCTCCTGGACCTTGGGGGTTCCCCCA

CTACAAGGACCGTCGGGGGCTGACCCCTCTGTTCACACGCCATGGTGGGTGGTGACCCCGATGCTGC
 GAGCTGCTCCTGTTC AACAGGGCCAGCTGGGCATAGCTGATGAGAACGGCTGGCAGGAAATCCACCAGG
 CCTGCCAGCGGGGTCACTCTCAGCACCTGGAGCATCTGCTTTTCTACGGGGCTGAGCCTGGAGCCCAGAA
 CGCCTCGGGGAACACGGCTCTGCACATCTGCGCCCTCTACAACAAGGAGACCTGTGCCAGGATCCTCCTG
 TATCGAGGTGCCGACAAGGATGTGAAGAACAACAACGGACAGACCCCTTCCAGGTGGCAGTGATTGCTG
 GGAATTTTGAGCTGGGGGAGCTGATCCGAAACCACCGAGAACAGGATGTGGTGCCCTTCCAGGAGTCCCC
 CAAGTACGCGGCCCGGCGACGGGGGCCCCAGGCACAGGGCTGACGGTGCCCCCGCGCTGCTGCGGGCC
 AACAGTGACACACGATGGCGCTGCCCGACTGGATGGTGTCTCCGCCCGGGGGCCGCGTCTCTGGGG
 CCCCTGGCCCTACCTCAGGGTCCCAGGGCCAGTCGCAGCCCTCGGCCCCACCACCAAGCTCAGCAGCGG
 GACCTCCGAAGTGCCAGCAGCCCCGGGGTGCCAGGGCCCGCTCTCCATCCCAGGGAGGCACCTTGAG
 GACGCCAAGAGGCAGCCCCGAGGCCGGCCAGCTCCAGCGGGACACCCGGGAAGGGCCAGCCGGGGGCA
 CGGGGGGCTCAGGGGGCCCCGGGGCTCCCTGGGCAGCCGCGGAGGCGGAGGAAGCTCTACTCAGCGGT
 ACCCGGACGCTCCTTCATGGCTGTGAAGTCTACCAGGCCAAGCCGAGGGGGAGATCTCCCTGAGCAAG
 GGCGAGAAGATCAAAGTACTTAGCATCGGGGAAGGAGGCTTCTGGGAAGGCCAGGTCAAAGGTCTGTGTG
 GCTGGTTCCCTCTGACTGCCCTGGAAGAAGTGGCGAATCGCTCTCAGGAGAGCAAGCAAGAAAGCCGAG
 TGACAAGGCAAGAGACTCTTCCGGCATTATACCGTGGGCTCCTACGACAGCTTTGATGCCCCAAGCTTA
 ATGGATGGGATTGGCCAGGGAGCGATTACATCATTAAAGGAGAAGACAGTCTTGCTGCAGAAGAAGGACA
 GTGAGGGGTTTGGGTTTCGTGCTCCGGGGGGCCAAGGCGCAGACCCCATCGAGGAGTTACCCCCACCCC
 GGCTTCCCGGCGCTGCAGTACCTGGAGTCGGTGGCAGGGGTGGCGTGCCATGGCGAGCTGGAGCTGCGA
 ATGGGAGACTTCTCATCGAGGTGAACGGGCAGAATGTGGTGAAGGTGCGCCACCGACAGGTGGTGAACA
 TGATCCGCCCAAGGGGGCAACACGCTGATGGTGAAGGTGGTGAAGGTGCGCCACCGACAGGTGGTGAACA
 GGCAGTGACAAGAAAGCACCCAGCAGGCCAAGCGGCTGCCGCCCCAACCATCTCCCTGCGTTCCAAA
 TCTATGACCTCAGAGCTGGAGGAGATGGAGTACGAGCAGCAGCCGGCGCCGGTGCCAGCATGGAGAAAA
 AGCGGACCGTGTATCAGATGGCTCTCAACAACTGGACGAAATCCTGGCCGAGCTCAACAGACCATCAG
 TGCAAGCGAAAGCCCTGGTCCCGGTGGCTCGCGTCCCTGGGCAACACCGACCCAAAGGTTTCTTTGGC
 ACTGAGTCGAGCTTCGATCCCCACCACCGTCCCCAGCCAAGTTACGAGCGTCTCTTTCTTCTGCTCCAG
 GACCTGGGTTGATGCTCCGGCAAAAATCTATCGGTGCGGCAGAGATGACAGACCTTACCTAGCACCCCC
 AGCCATGAAATTCAGCCGAGCCTGTCTGTGCTGGTTCCGAGGACATTCCCCCGCCACCCACCACGTAC
 CCACCGGAGCCTCCCTACAGCACACCTACAGTCCCCCTCCTCCTCAGGGCGCTCACCCCTCCCTTCGGG
 GAGGGCCCTTCAACCTGGCTCTGGTGGCCCCCTCCCCGCTCCTCCCCTGCAACCTTTGACGGGCCCTC
 CCTCCCGACACTCGCGTGGGAGCCGCGAGAAGAGCCTGTACCACAGTGGGGCCCTGCCCGCGCCAC
 CACCACCGCCCCACCACCCAGCCACCGCCCCCGCTCAGCCCCACCACCACCGCCACCCCC
 CTCATCTCCCGAGATGGAGACAGGCGGCTCTCCCGACGACCTTCCACCCGCTTGGCTCTGGGGCCCCA
 GCCCAGCCTGCGAGGCTGGAGGGGCGGCGGCCAGCCCGACCCGGGGGCCCCGTCCCCATCGCACCCAC
 GGCAGCGCGGGCGGGGCGGCGGCTCCTCCAGGGCCCGGCTCTACGCTATATCCAGCTGCCCCCGCGGG
 CGGCCAGCGCAGCCATGTACGTGCCCGCCGCTCGGGCCGCGGCCGCAAGGGCCCGTGGTCAAGCAGAC
 CAAGGTGGAAGGCGAGCCCCAGAAGGGCGCGGCTCCTCGCCCGCGCCGTGCGCCCTCGTCCCCGGCTCC
 CCGCAGCCGCGCCGCGCGCTGGCCGCGCCCTCGGAGAAGAACTCCATCCCCATCCCCACCATCATCATCA
 AGCCCCCGTCCACCAAGTAGCAGCGGCCGAGCAGCCAGGGCAGCAGCACCGAGGCGGAGCCCCCACCCA
 GCCGAGCCACGGGAGGCGGCGGCGGCGGCTCCTCGCCAGCCCCCGCCCGGCCATGTACCCGCTG
 CCCCCGTCCCCCTCGCCCGTGCCACCCCGCTCGCCAGCGGCCCGCCACGCTGGACTTCACGAGCC
 AGTTCCGGGGCGCCCTGGTGGGGCGGCCCCGAGGGAGGGGGCTGGCAGAATGAGGCGCGCGGCGCTC
 CACGCTGTTCTGTCCACCGACGCGGGGACGAGGACGGCGGGACGGCGGGCTGGGCACAGGGGCGGCC
 CCGGGCCCGCGGCTGCGCCACTCCAAATCCATCGACGAGGGCATGTTCTCGCCGAGCCCTACCTCCGAC
 TGGAGTCTGCGGGCAGCGGCGGGCTACGGCGGCTACGGGGCCGGTAGCCGAGCCTACGGGGGTGGCGG
 GGGCAGCAGCGCTTCACAGCTTCTTGCCCCCGGACCCCTGGTGACCCGCTGACCGCAAGGCCCTG
 GATCCCGCTCCCCGTGGGGTGGCCCTGGCCGCCCGCGAGCGAGCGCTGAAGGAGTCTCGGAGGGCG
 GCGGGGCCCCCAGCCGCTCCAGGCCCCATCGCCCGCTACGAGGCCCGCGCCGCCACCCCGCACCA
 CCACTCGCCCCACGCCACCACGAGCCAGTGTGCGTCTCTGGGGGGCTCCCCGCGGACCTTGCAGCGC
 CGGGAGCTGGGGTACAGGGCCGGGTGGGACGCCAGGAGAAGTCCCTTCCCGCCAGCCCGCCCGCGCC
 GGCGTTCCCTGCTACACCGCTTCCGCCCCACCGCTCCCGGGTGGGGCCCCCTCTGCTGCAGCTGGGGAC
 GGAGCCCCCGGCCCGCACCCCGAGTAAGCAAGCCCTGGAGGTCCGACGCCCGCAAGAACCCGAGCGG
 CTGCCGCTGCAGTGGGTTCTTTGAAAATGCCAGCCCCGGGCCCTGTGACGAGCGGAAGGGGTCCCC
 CCTCGGAGGACGGGCGGGGTCGCCCGCCAGCCACGCGGTCCGTCGCCCTTCCCGACCTCCCC
 GAGGGCCAGCGAAGAGAACGGGCTGCCCTGCTGGTCTGCGGCTTCCCGCCCCCTCGGTGGATGTGGAA
 GATGGCGAATTCCTTTTCGTGGAACCGCTGCCTCCGCTCTGGAATTCCTCAACAGCTTCGAAAAGCCAG
 AGTCGCCCTCACGCTGGGCTCCCCACCCGCTGCCCGACACCTGCCCTGCCACCCCGTTACCCCC

TGTGCCACCCCGGCTGTGGCCGACGCCCTCCACCCCTGGACTCCACCGCATCCAGCCTGACATCCTAT
 GACAGCGAGGTGGCCACCCTGACCCAGGGGGCTCCGCCGCTCCTGGGGACCCCATCCACCAGGCCCGC
 CTGCCCCAGCAGCACCGGCTCCCGCTGCCCCACAGCCTGGCCCGGACCCCTCCGCTGGCACGGATTCTGG
 CATCGAGGAGGTGGACAGTCGGAGCAGCAGTGACCACTGGAGACCATCAGCAGCGCTCCACGCTG
 AGCAGCCTATCTGCCGAAGGTGGTGGCAGCGCAGGGGGTGGGGCGGGGCTGGGGCCGGTGTGGCCAGTG
 GGCCGGAGCTTCTGGACACCTATGTGGCTTACCTGGACGGCCAGGCCTTTGGGGGACAGTACTCCCGG
 CCCGCCATACCCCTCCTCAGCTCATGACTCCCTCTAAGCTCCGGGGCCGGGCGCTAGGAGCCAGCGGAGGC
 CTGCGGCTGGGCCCAGCGGGGACTCCGAGACCTGTATACCCCAACAGCCCAACCGTCTCGGTGACAG
 GGCTTGAACCGATGGGCTGCTGGCCCTGCGTGCTTGTTCAGGACCCCCACGGCAGGCGTGGCGGGGG
 TCCGGTGGCTGTAGAGCCAGAAGTCCACCGGTGCCCTTGCCGACGGCCTCCTGTCTGCCCCGGAAGCTG
 CTGCCCTGGGAGGAGGGCCCGGGCCACCGCCACCACCTCTGCCCGGGCCCTTGCCAGCCTCAGGCCT
 CAGCCTTGCCACAGTAAAGCCAGCATCATCAGTGAATCAGCTCCAAGCTTCAGCAGTTTGGGGGCTC
 CTCGGCAGCTGGCGGCGCTCTGCCCTGGGCCGAGGAGGCAGTGGGGGAGGCGGAGACAGCCACCACGGG
 GGAGCCAGCTATGTCCCCGAGAGGACCTCCTCCCTGCAGCGGCAGAGACTCTCCGACGACTCCAGTCTT
 CACTCCTCTCCAAGCTGTGACGAGCCTGTTTCAGAACTGGCCCAAACACCTCTGCCGCCACTCCCCAC
 CGGAACAGGGGTCTCCCTACAGCCGCTGCGGCCCCAGGGGCCACCTCACCCTCAGCCTCCTCCTCTCC
 ACGTCCACCCGCCACCTCCAGGGCGTGGAGTTTCAGATGCGGGCCCCCTCTGCTCCGCCGGGCCCCAGCC
 CCTCGCTGCTGCCGCCCTCGGAGCACAAAGTCAGCCCTGCGCCAGGCCCTCGTCCCTGCCCATCTTGCC
 TTCCGGACCCCTCTACCCAGGCCTCTTTGACATCCGTGGCTCCCAACTGGAGGGGACAGGAGCCTCGGCT
 GACCCCTTGCCCGCCAGTCTTTGTGCCGCCACACCGGGATATCCGGGGGGCTCGGGGGAGCCTTGTGAC
 GGCCCTCGCGCTCCCTCTCACCGACCCGCTGCTCTCGCTGCCCGGACAAGCCGTTTGGCGCTAAACC
 TCTGGGGTCTGGACCAAGTTCGACGTGGCTGATTGGCTGGAGTGGCTGGGTTTGGCGGAGCACCGAGCC
 CAGTTCCTGGACCACGAGATCGATGGCTCCACCTGCCCGCCTTGACCAAGGAGGACTACGTGATCTAG
 GTGTGACCAGGGTGGGCCACCGCATGAACATGACCGGGCTCTCAAATCTTCTTGAGAGGTGATGGCT
 GGCTTGACGGACAGCCCGTCCACAGAACTCTTGAGCCTGCTGGCCTCTTGACCTCTGACCCCTGACT
 GTCATTCTCTCCCGGGCCAGGGAATCTGTTCAAACCTGCGCCCTGCCCTCATCTTCCAAGGCC

>gi|19743793|ref|NM_012309.1| Homo sapiens SH3 and multiple ankyrin
 repeat domains 2 (SHANK2), transcript variant 1, mRNA
 ATGATGATGAACGTCCCCGGCGGAGGCGGCCGCGGTGATGATGACGGGCTACAATAATGGTCGCTGTG
 CCCGAATTCTCTCTACAGTGACTGCATTATTGAGGAGAAGACGGTGGTCTGTCAGAAAAAAGACAATGA
 GGGCTTTGGATTCTGCTGCTTCGAGGGGCCAAAGCTGACACACCCATTGAAGAATTCACACCAACACCGGT
 TTCCAGCCCTACAGTACCTGGAGTCCGTGGATGAAGGTGGGGTGGCGTGGCAAGCCGGACTAAGGACCG
 GGGACTTCTTGATTGAGGTAAACAATGAGAATGTTGTCAAAGTCGGCCACAGGCAGGTGGTGAACATGAT
 CCGGCAGGGAGGGAATCACCTGGTCTTAAGGTGGTCACGGTGACCAGGAATCTGGACCCCGACGACACC
 GCCAGGAAGAAAGCTCCCCGCCTCCAAAGCGGGCACCAGCACAGCCCTCACCTGCGCTCCAAGTCCA
 TGACCTCGGAGCTGGAGGAGCTCGTGGATAAAGATAAACCAGGAGATAGTCCCGGCTCCAAGCCCTC
 CCGCGCTGCTGAGAACATGGCTGTGGAACCGAGGGTGGCGACCATCAAGCAGCGGCCACGACCGGTGC
 TTCCCGCGGGCTCAGACATGAATCTGTGTACGAACGCCAAGGAATCGCCGTGATGACGCCCACTGTTC
 CTGGGAGCCCAAAGCCCCGTTTCTGGGCATCCCTCGAGGTACGATGCGAAGGCAGAAATCAATAGACAG
 CAGAATCTTTCTATCAGGAATAACAGAGGAAGAGCGGCAAGTTCTGGCTCCTCCAATGCTGAAGTTCACC
 AGAAGCCTGTCTATGCCGGACACCTCTGAGGACATCCCCCTCCACCGCAGTCTGTGCCCCCGTCCCCAC
 CACCACCTTCCCAACCACTTACAACGTGCCCCAGTCCCCCAACTCCAAGAGTCTACGGGACGATTAAGCC
 TGGCTTCAATCAGAATCTTGCCGCCAAGGTGTCCCCCGCCACCAGGTCCGACACCGTGGCCACCATGATG
 AGGGAGAAGGGGATGTACTTCAGGAGAGAGCTGGACCGCTACTCCTTGGACTCTGAAGACCTCTACAGTC
 GGAATGCCGGCCCGAAGCCAACTTCCGCAACAAGAGAGGCCAGATGCCAGAAAACCCATACTCAGAGGT
 GGGGAAGATCGCCAGCAAAGCCGTCTACGTCCCCGCCAAGCCCGCCAGGCGGAAGGGGATGCTGGTGAAG
 CAGTCCAACGTGGAGGACAGCCCCGAGAAGACGTCTCATCCCTATCCCGACCATCATCGTGAAGGAGC
 CGTCCACCAGCAGCAGCGGCAAGAGCAGCCAGGCGAGCAGCATGGAGATCGACCCCCAGGCCCCGGAGCC
 ACCGAGCCAGCTGCGGCCCTGACGAAAGCTGACCGTCAGCAGCCCCCTTGCCCGCCGATCGCCGGAGCC
 GTCCGCGACCGTGAGAAGCGGTGGAAGCCAGGAGGAATCCCCGGCTTCTCTCCACAGACCTGGGGG
 ATGAGGATGTGGGCTTGGGGCCACCCGCCCCAGGACCGGGCTCCATGTTCCTCCGAGGAGGGGATTT
 TGTGACGAGGACAGCGCTGAGCAGCTGTATCCCCCATGCGGAGTGCCACGCCCCAGGAGGCCGAAAC
 CATTTCTGGGTGGCGCCGAGGCCAGTGTCTCCGGTGGAGCTGGGAGGCGCTGAATTCACGTCCAAAG
 CCCAGGGGCCGAGAGCAGCCAGCAGTGGCTCCCGCAGCAGCGGCACAGCCGGCCCCGGGAATTTATGT
 CCACCACTCACAGGGCGGTGCTTGATCCAGCTCCCCGCTGGCCCTGGCACTCTCCGCAAGGGACCGA
 GCCATGAAGGAGTCTCAACAGGACCCAAAGGGGAGGCCCCCAAGGCCGACCTCAACAAACCTCTTTACA

TTGATACCAAAATGCGGCCAGCCTGGATGCCGGCTTCCCTACGGTCACCAGGCAGAACACCCGGGGACC
 CCTGAGGCGGCAGGAGACGGAGAACAAGTACGAGACCGACCTGGGCGGAGACCGGAAAGGCGATGACAAG
 AAGAACATGCTGATCGACATCATGGACACGTCCCAGCAGAAGTCGGCTGGCCTGCTGATGGTGCACACCG
 TGGACGCCACTAAGCTGGACAACGCCCTGCAGGAAGAGGACGAGAAGGCAGAGGTGGAGATGAAGCCAGA
 CAGCTCGCCGTCGAGGTGCCAGAAGGTGTTTCCGAAACCGAAGGTGCTTTACAGATCTCCGCTGCCCCC
 GAGCCCACCACCGTGCCCGGCAGAACCATCGTCGCGGTGGGCTCCATGGAAGAGGCGGTGATTTTGCCAT
 TCCGCTCCCTCCCTCCCCCTCTGGCATCCGTGGACTTGGATGAGGATTTTATTTTACAGAGCCATTGCC
 TCCTCCCTGGAAATTTGCAAAATAGTTTGTATATCCCCGATGACCGGGCAGCTTCTGTCCCGGCTCTCTCA
 GACTTAGTGAAGCAGAAGAAAAGCGACACCCCTCAGTCCCTTTCGTTGAACCTCCAGCCAACCAACCACT
 CTGCAGACAGCAAGAAGCCAGCCAGTCTTTCAAACGTCTGCCTGCCTCATTCCTGCCACCCCTGAAAG
 CTTTGACGCCGTCGCGGACTCTGGGATCGAGGAGGTGGACAGCCGGAGTAGCAGCGACCACCACCTCGAG
 ACGACCAGCACTATCTCCACCGTGTCTAGCATCTCCACCTGTCTTCCGAAGGTGGAGAGAATGTGGACA
 CCTGCACAGTCTATGCAGATGGGCAAGCATTTATGGTTGACAAACCCCAAGTACCTCCTAAGCCAAAAAT
 GAAGCCCATCATTCACAAAAGCAATGCACTTTATCAAGACGCGCTCGTGGAAAGAAGATGTAGATAGCTTT
 GTTATCCCCCGCCCGCTCCCCCGCCCCGCCGGGCAGTCCCGAGCCTGGGATGGCCAAGGTTCTCCAGC
 CAAGGACCTCCAAGTTGTGGGGCGACGTCACAGAGATCAAAAGCCCGATTCTCTCAGGCCCAAAGGCAAA
 CGTTATTAGTGAATTGAACCTCTATCCTACAGCAAATGAACCGAGAGAAATTTGGCAAAGCCGGGGGAAGGA
 CTGGATTACCAATGGGAGCCAAGTCCGCCAGCCTCGCTCCAAGAAGCCGGAGATCATGAGCACCATCT
 CAGGTACACGGAGCAGCAGCGTCACCTTCACTGTTCCGCCCGGCACCTCCAGCCCATCACCCTGCAGAG
 CCGGCCCCCGACTATGAAAGCAGGACCTCAGGAACAAGACGTGCCCAAGCCCTGTGGTCTCGCCAACA
 GAGATGAACAAAGAGACCCCTGCCCGCCCCCTGTCTGTGCCACCGCTCTCTTCTCCCGCTCTCTCAG
 ATGTCTTTAGCCTTCCAAGCCAGCCCCCTTCTGGGGATCTATTTGGCTTGAACCCAGCGGGACGCAGTAG
 GTCGCCATCCCCCTCGATACTGCAACAGCCAATCTCAAATAAGCCTTTTACAACCTAAACCTGTCCACCTG
 TGGACTAAACCAGATGTGGCCGATTGGCTGGAAAGTCTAAACTTGGGTGAACATAAAGAGGCCCTTCATGG
 ACAATGAGATCGATGGCAGTCACTTACCAAACCTGCAGAAGGAGGACCTCATCGATCTTGGGGTAACTCG
 AGTCGGGCACAGAATGAACATAGAAAGGGCTTTGAAACAGCTGCTGGACAGATAAGGACGGCTGCTCTCC
 ACCTCGCAGACTGCTCTTGTATAAGTAGAGATGGGCTCGTGCTGAAACATCTGAATGCCAAGCGAAGTC
 TGTGAGCATCAACCCCACTCCATGGGTGTGTCTCTGGTACCCAAAGAAATACAGAGTTGTGTCCACAAC
 ATGGCTGGGTCTTCAGACCCCTGGCTCACCATGTGGGTGTCTTGGGCAGTTTCTATCACACATGGGACAA
 GGGGAGGGAGTTTCTAACATGGAAGAAAGATTCTCAGCCTGCCGCCAGCATGCAGGTGGCCTCGCTTT
 GCCGGTCCGAGAGGCTCCCCGTCAATTTGCACGGGATCCTAGCTCTTGTAGGCAGACACCAGTCGACT
 CTAGATACCTCCTGAGACCTCCGTCTCTGCTTTCCGGGCAGCTCTCACCACCCAGGCCCGGCATGAG
 GCCTTTCTCAGTCTGTGGCCTCTCAGAGGACACCTGATGCTCAGCTGCCCTCTTTCTCTCTGCACTTG
 GCTTGCAGTGAGATGCTCCAGATGCATTTGTCCAGTGCCCCATCATGGGCCGTAAGGCAGAGAACTT
 TTTCTACACAGATTCTTTTCCCCATCTCCTCCTGTGGTTTGCATCCATGGCTCTTTGGCCATGAGGTTT
 CTGGCAGTGTGGGAGTTTGGATGGGATCGTGCCAGCTTTGCTTAGCTTTCTTTATTTCTGCAAATCTG
 TTAGCATAATTTCAAGGTGGCCAAGCAGATGTCACATGGAGTTAGTCAAAGCACAAAGTCACGATTCCAC
 AATGGAGGGGAGACCTGGCCAAGGGAGCCAGCCAGCGTGCAACTGCCCAAGCTCCAGGTCTCCAGGACAA
 GAGCAGTTGTCTGCCATGAGCACCCATCCAGGATGGAGAATAAGGGCTTCTCTGCCCTCTCAGAATTCTTT
 TTAATTGAAGATGTCTTGAGCTCTGCAAGATCAGAGCAGGTGAGCATCCACTTTGACATGAAGGACAAG
 AAGACGCATGGCTCATGGCGGCACATGCGGCTGCCAGTGAGACAGCGTCTCCTCTGGGAGCTGGGCGGG
 CACAGCATCTCAGTCTGTGCCCAGCCAAGGGTGAGCATCTGCTGAGACAGTCTTTTGCTCTCGGA
 GGCCAGGGAAGATGGTACTTAGAGGCTTTTCCCTATCGCTCTGGGTGTCTAGGAATCCCACAGCTTGT
 CTTAACAGTACAACAGCTTCTTTGAGGACCCAGTGGGTATGGAGTATAGACAGAACCAGGGTTGAGAAC
 AGAAGGTGGGCGGCAGGATCAGAGTGAAAGCAGAGGCGTGAGGAGAGGAAAGCAGGGAGGTCTCTGGGC
 TGCCAGGTACGCTCTCTGGCAAGGCTTCTTTGAGCCCCGCCCTTTCTTTCCCCGAGTCCCTCCACCC
 CATAACAATACCTCGAATTTCCAAAAGAGGTCACCAGATGCACATGGGCCGCAAAACACACAGTCAGGCT
 TCCAGCACATTTCCCCCATTTGGAGGATCTCGAATGTGAGGTTTTTGGTTTTATTATTATTTTCAAGAC
 TAGCTCAGCCCATCTCTAATTATAAAACATGGTTTTGTTTTTTTTTTCTTTTCTTTGATTAGGTC
 TGGAACAGCTCTAGAATGAACACATAAAATTTAGCAATTTAAATCTTTCTTTACTGCAAGTTTAAATAG
 TTGTACAGATAGTTTATAAGCACAATATTTTAAAGAAAAAAAGTGGCTGGTCTACTAGGCAGCCTTTGTG
 CCACTTCAGTGTAGAAAGTTAAAGAAAAAAACCTTTGTGATTTAATAATACTATTTCTGTGGAATAA
 TTATAAAAGTATGACCTTTTAAATCAACCTTATTTGGATGCATCTGAACCAGCAGAGCTGTGTTATATT
 TTCTATCTTTGTAGAACTTCGTCAATTGAAGGACAATTTCTTCAAAGTGTTTACAATTCATAATTCAGCA
 GTTTCTTCCAAAAACAAAAACAAACACACACACACACGCGCTTTTCCAGTCACACACCCCTGATGTT
 GGAACCAAGTTTTTGGACCTTCTGTTCAAAACCTTTTGAGGTCAATCTTTGTATTTGAAATGATCCAA
 TCCAACCTGAAGTCAATTGAATATTAAGGCGCTTTACTTCCGTGTGCTTTCAGTTTTTCCATCATGAGAT

GAATGAGCATTACTCTAGATAAAATTTCAAGACAGGATACTACAGGTGGCCTGCTGAGGCTGCCCCATATT
TTAGAAAATGTAAAAATGGTGGTTTGGCCATTAATTTGTCTTCCATTTGATGATACCGCAAATTCGGTG
AGTCCATTCCCTTTGGCATGGCACTTTCCCTGGGCCTACAGTTGGTATTACCTCTGTGCTCAGTGCCAGGC
AAAACACTAGCTCAAAGGAGAGTCAAGGAAACCGCTGGCGAGACGATAACCAGTCGAAACTCGTGACTTCG
GTTTGTGTGAACCTTTGGCAGCCAGTTGGTGAGGGCCAGATGTTATTCCCTTTCTTAAAGATACTCCAAGCC
ACATGCCACTAACCACAAGCAAGCTGGCTGCAAGACTAAAGAGCTGATAACATAGTTTATTTTACACTG
TCTTATTATAGAGAAGTAATAGACCTATCAGAACCCTGCCTGACCAACAATAAACACATGTGTGCCAAGA
TGAATCGGTCTCTATCTCTATCTGCTTATTTTGGTATGAAAGCAATAGTTCTCTCATTCAAATCACCAC
CACTGTTCTTCCCCCTTTGGGACATGTTTAGACGAGGGCCCTATTCCATGCCCCCTCTTTAATGGTGGAACAA
ATGTTAAACTGCTCATCTAAAGATCATGTTGATATTATTCCAGGTTTTAAGATCAACTTTTGTGTACATAC
TGTAATTTAAATAAACGTCATTTACATGCCTAGTTTCTGTAATATTGTGTATACAAAACCCAAATCTCTC
AAAATGTAAATTATGTATACCTGCCAAGATACCTTTTCCAGGGTGTCTGCGCACATTTTAAGTTAATTCA
CATAATATAAAAAATTACTCAATGTGACTGTTGATTGTGCTGAACCTTTACATATCACAAAGTGAATTATTTG
TGATACTTTAGTTAATAAAATGGTAAATTTTTTCTCAGTTATTGAACAAGCAAGCATTATCCAGTTGAT
CTGGCAATGACTTTTTTGTGTGTGGGCCACAATATTGATTTTCCCATTAACAATTTTTTTTTTGTTTTTTAA
ATACTAATATGTTTCACACTATAGTTTGTGTAACAACACGTGTTTCGCATTATCTATGTTGCTGTTACTTT
TGTGCTTTTATTCTTTTTTAGACTTTATAAAAAAAAAAAAAAGCTCCTGTAATTTGCACCTTTCTCCCAATC
CTTAAATCTCTTGTATGGCAACCAAAATACTGTAAAAAAATAATATACATTATGCACTAAGGTTGTGGT
TCTGATTGCAAAACAAACAGCTGAACACTGTCTGAATTTAAACAAAAAGCTGCCGCACTTGAACCTTAATGTGTA
GATTATCTCAGGCATTGTGGCCAGCTCTGCCCTCTCTAAAACTGACCAGAAAAATCTCTCTCATCGAGTAA
ACAGGCTCCTGTCACTGAGCTAATCTGCCCTTGGTTCATTTCCTTATTCTCAATTTATCAATGGATACGT
GCATGTTATTTCAGAATTATGCAAAACGTCAAAATCTGCTTCTGTGACCGCTGCTATAGGCGTGGAGCTG
AGGCTCGGCTTTTCTTTTTGTCTGGGTGGAAGCAGCGGTGCCGCGGAGGGCCAGCCAGATCCGAGCCCT
TCCCTTAGGGTCCAGTCTCCCCACACCCAGCAGGGTGTCTTCTAGCCATAAGGCCAAGGGAGTGGCAGA
ACTGGGCCCGCTCTCTGGTTGACAAGCAAACCACATGCTAAGGCTTGGAGCAAGAGAGAATTTGTGTCTA
TTGGCAAAGAACTAAGCCAGGAAGACATGGGCCATCCCTCCGCTTTAGGGAAGCATATTTTAAACCTAAA
CGTTGAACCTCTTCTTTGGCCTCACCAGTGA AAACTTGTGTGCTTTAGTTCCCTAAAGTTCTTCTACTTT
GGCACATTCCCCACTTGAGCAGCAGCCTCTATGCTTCCACGTTCAGGA AAAATCCAGTCCCTCATATCTT
TTGTAGTTACCCCTCAAGCTCTCCCGCTTACCATTCCAAATAGTTTCTCCCAAACTTTGGCACCCTCCCTAG
ACTTTGCTTCTCCCAATGGTTTCTTCCAGACCATTPTTCTAGATGAATATATTCTGTTTACCTTACTAGGAAA
ATTATTGGAAGATTTTTTCTTTTACTTGAAATTGGAGGCATTTAATAACTGGCGAACTGGAATGTGTTT
CTGTATTGTAGACAACCATGAACCCATGCAAGTAGGTGAACATTCCACAGTGGCTGGGTGACCACAGCA
GCTGCATGCAGACAGGACTGCCCGTGCTTTGTGGGAATCAGAGAATTTCCAAACTTGTTTCTCAGACTT
CCGCAGATCTCATCACTTTGATTTCTAATCCATGCTGTATTGGTGATTTTGTTTATCGTTTCTGTAACCT
GTTCTACATTCCACAGTCTTTACCGTTTTATGTTCAAAATTAACAACAATCCCTGTCCATTGATTCCACTC
TGGAACTCTTTGTTTATGCCAATTTTGAATTTTAAATACGAGCCTTCAAATAAACACAGAAAAGAAAAAA
AAAAAGGAAAAAAAATAAAGTGTGTGCCCTGGGTGTTTGAACCTGATCTGCATAAAACCATATTGTATT
CAAATGCTGTATCTTTTATGAAAGGATAAATATATCTTCAATGTATTAAATGTAGCTGATAAAATAGTTAT
AGTGCCTTGTTTTCTGAACTGAGAAGCCCATATCAATGCACAACGTAAAGAAAGGACGGTTCGAGCTATG
AAATGTTGAGGACACCTACACGTGATTCTGGAGCCTTTCTCATCATAAGACTGTGGAAAAATACACAGTCTG
TTTAAGTTAAAAATAAAAGTAAATTGA

```
>gi|37564791|ref|XM_037493.2| Homo sapiens SH3 and multiple ankyrin repeat domains 3 (SHANK3), mRNA
```

ATGCAGCTGAGCCGCGCCGCGCCGCGCCGCGCCGCGCCCTGCGGAGCCCCGGAGCCGCTGTCCCCCG
CGCCGGCCCCGGCCCCGCCCCCCCCCGCCCCCTCCCGCGCAGCGCGGCCGACGGGGCTCCGGCGGGGGG
GAAGGGGGGGCGCGGGCGCCGCGCGCGGAGCGGGGCCCGGGAAGCCCCGGCCCCCTGGGGCGGGGGCTTCG
GCCGCGAAGTGCCTGCGCCTGGACCCGGCCGCGCCCGTGTGGGCCCGCAAGCAGCGCGTGCTCTGCGCCC
TCAACCACAGCCTCCAGGACGCGCTCAACTATGGGCTTTTCCAGCCGCCCTCCCGGGGCGCGCCGGCA
GTTCTCGATGAGGAGCGGCTCTGCGAGGATACCCGCCCAACTGGACACGCCCCTGCCCTACCTGGAG
TTTCGATACAAGCGCGAGTTTATGCCAGAACTCATCGATGATAAGCAGTTTGCAAAGCTTCACACAA
AGGCGAACCCTGAAGAAGTTCATGGACTACGTCCAGCTGCATAGCACGGACAAGGTGGCACGCCCTGTTGGA
CAAGGGGCTGGACCCCAACTTCCATGACCTGACTCAGGAGAGTGCCCCCTGAGCCTCGCAGCCAGCTG
GACAACGCCACGGACCTGCTAAAGGTGCTGAAGAATGGTGGTGCCCACTGGACTTCCGCACTCGCGATG
GGCTCACTGCCGTGCACGTGCGCACACGCCAGCGGAATGCGGCAGCAGCTGACGACCCCTGCTGGACCTGG
GGCTTCACTGACTACAAGGACAGCCGCGGCTTGACACCCCTTACCACAGCGCCTGGGGGGTGGGAT
CGCCTCTGCTGTGAGCTGCTTCTCCACGACACCGCTCAGCTGGGGATACCGACAGAGAATGGCTGGCAGG

AGATCCACCAGGCTGCCGCTTTGGGCACGTGCAGCATCTGGAGCACCTGCTGTTCTATGGGGCAGACAT
 GGGGGCCAGAACGCCCTCGGGGAACACAGCCCTGCACATCTGTGCCCTCTACAACCAGGAGAGCTGTGCT
 CGTGTCTCTGCTCTTCCGTGGAGCTAACAGGGATGTCCGCAACTACAACAGCCAGACAGCCTTCCAGGTGG
 CCATCATCGCAGGGAACCTTTGAGCTTGCAGAGGTTATCAAGACCCACAAAGACTCGGATGTTGTACCATT
 CAGGGAACCCCCAGCTATGCGAAGCGGCGGCGACTGGCTGGCCCCAGTGGCTTGGCATCCCCTCGGCCT
 CTGCAGCGCTCAGCCAGCGATATCAACCTGAAGGGGGAGGCACAGCCAGCAGCTTCTCTGGACCTCGC
 TGAGAAGCCTCCCCACCAGCTGCTGCTCCAGCGGCTGCAAGAGGAGAAAGATCGTGACCGGGATGCCGA
 CCAGGAGAGCAACATCAGTGGCCCTTTAGCAGGCAGGGCCGGCCAAAGCAAGATCAGCGCCGTCCCCGGC
 CGCAAGTTTCATCGCCGTGAAGGCGCACAGCCCGCAGGTGCTCAGCATTGGGGAGGGCGGTTTCTGGGAGG
 GAACCGTGAAAGGCCGCACGGGCTGGTTCCCGGCCGACTGCGTGGAGGAAGTGCAGATGAGGCAGCATGA
 CACACGGCTGAAACGCGGGAGGACCGGACGAAGCGGCTCTTTCGGCACTACACAGTGGGCTCCTACGAC
 AGCCTCACCTCACACAGCGATTATGTCAATTGATGACAAAGTGGCTGTCTGCAGAAACGGGACCACGAGG
 GCTTTGGTTTTGTGCTCCGGGGAGCCAAAGCAGAGACCCCCATCGAGGAGTTACGCCACGCCAGCCTT
 CCGGGCGCTGCAGTATCTCGAGTCGGTGGACGTGGAGGGTGTGGCCTGGAGGGCCGGGCTGCGCACGGGA
 GACTTCTCATCGAGGTGAACGGGGTGAACGTGGTGAAGGTCGGACACAAGCAGGTGGTGGCTCTGATTC
 GCCAGGGTGGCAACCGCCTCGTCATGAAGGTTGTGCTGTGACAAGGAAGCCAGAAGAGGACGGGGCTCG
 GCGCAGAGCCATGAAGAAGTTTGCCTCCAGCCGTAACCTGAACAAGATCCTGGCACAGTGCAGACTCATCC
 TCGCGGGAGTACGAGGAGATCCAGGCGGTGGAGCGCAAGTGGCACCCTGCACCTGGCCACGCCGCGCCGCC
 TGCTGCTGGACAGGAGGTCCAAGGCCCTCCCTCTTCTTTGAGCCCCACCGCCCCCAAGAGGGCCCCAG
 CACCACACTGACCTGCGCTCCAAGTCCATGACAGCTGAGCTCGAGGAACCTTGAGAAGCTGGACGATG
 CTGGCAGCCGCCGCGAGGCCAACGCTGCGGCCAGACATCGCAGACGCAGACTCCAGAGCCGCCACCGTCA
 AACAGAGGCCACCAGTCGGAGGATCACACCCGCCGAGATTAGCTCATTGTTGAACGCCAGGGCCTCCC
 AGGCCCAGAGAAGCTGCCGGGCTCCTTGCGGAAGGGGATTCACGGACCAAGTCTGTAGGGGAGGACGAG
 AAGCTGGCGTCCCTGTGGAAGGGCGCTTCCCGCGGAGCACCTCGATGCAAGACCCGGTGCAGGAGGTC
 GCGGCATCCCGCCCCCGCCGAGACCGCGCCGCTTCCCCGCCCGCGCCCTACTACTTCGACTCGGGGCC
 GCCCCCGGCTTCTCGCCGCCCGCCCCCGCGGGCCGCGCCTACGACACGGTGCCTCCAGCTTCAAGCCC
 GGCTTGGAGGCGCGCTGGGCGCGGGCGCTGCCGGCTGTACGAGCCGGGCGCGGCCCTCGGCCCGCTGC
 CGTATCCCGAGCGGCAGAAGCGCGCGCTCCATGATCATCTGCAGGACTCGGCGCCCGAGTCGGGCGA
 CGCCCCCTCGACCCCGCCCGCGGCCACCCCGCCGAGCGACCAAGCGCCGGCCCGGCCCGCCCGGCC
 GACAGCCCCCTACGCCAACCTGGGCGCTTTCAGCGCCAGCCTCTTCGCTCCGTCCAAGCCGAGCGCCGA
 AGAGCCCCCTGGTGAAGCAGCTGCAGGTGGAGGACGCGAGGAGCGCGGCCCTGGCCGTGGGCGCC
 CGGTCCCGGGCGGCGAGCTTCGCGCGCGAGCCTCCCGACCAACCGCGGTCCGCGCCGGGTGGCCTC
 GACTACGGCGCGGGCGATGGCCCGGGCTCGCGTTCGGCGGCCCGGGCCCGGCCAAGGACCGCGGCTGG
 AGGAGCGGCGCGCTCCACTGTGTTCTGTCCGTGGGGGCCATCGAGGGCAGCGCCCCCGCGCGGATCT
 GCCATCCCTACAGCCCTCCCGCTCCATCGACGAGCGCTTCTGGGGACCGGCCACCAGCGCGCGGAC
 CTGCTGCTGCCCCCCCCGGTGTCTGCCCTGAAGCGTTGGTACGCGGCCGAGCCTGGGGCCCTCGGGTT
 CCACCTTCATCCACCCACTCACCGGCAACCCCTGGACCCAGCTCACCCCTGGCCCTTGCCCTGGCTGC
 CCGAGAGCGAGCTCTGGCCTCCAGGCGCCTCCCGGTCCCCCACACCCGTGCACAGTCCCGACGCCGAC
 CGCCCCGACCCCTGTTTGTGGATGTACAGGCCCGGACCCAGAGCGAGGGTCCCTGGCTTCCCCGGCTT
 TCTCCCCACGGAGCCAGCCTGGATTCTGTGCTGCTCGCAGGAGGCGAGAGAAGTCCCCCGGGAGGA
 GCGGAAGTCACCCGAGGACAAGAAGTCCATGATCTCAGCGTCTTGACACATCCCTGCAGCGGCCAGCT
 GGCTTCATCGTTGTGACGCCACCAGCAACGGGCAGGAGCCAGCAGGCTGGGGGGGGCCGAAGAGGAGC
 GCGGGGACCCCGGAGTTGGCCCCGGCCCCCATGCAAGTCAAGCGGCTGTGGCAGAGCCCCTGCCAGCCC
 CCGGGCCAGCCCCCTGGTGGCACCCCGGCAGACGCCGGGCCAGGCCAGGGCAGCTCAGAGGAAGAGCCA
 GAGCTGGTGTGTTGTGTAACCTGCCACCTGCCAGCTGTCTGTCAGCGATGAGGAGACCAGGGAGGAGC
 TGGCCCGAATTGGGTGGTGCCACCCCTGAAGAGTTTGCCAACGGGGTCTGTGTCACCCACTCGC
 TGGCCCGGGCCCCCTCGCCACACCGGTGCCAGCCGGCTCAGGGAAGCCCAGCAGTGAGCCACCCCT
 GCCCCTGAGTCTGCAGCCGACTCTGGGGTGGAGGAGGCTGACACACGAGCTCCAGCGACCCCCACCTGG
 AGACCACAAGCACCATCTCCACGGTGTCCAGCATGTCCACCTTGAGCTCGGAGAGCGGGGAACCTACTGA
 CCCCCACACCTCTTCTGCTGACGGACACACTTTTCTACTCGAGAAGCCACCAGTGCTTCCCAAGCCCAAG
 CTCAAGTCCCCGCTGGGGAAGGGGCGGTGACCTTCAGGGACCCGCTGCTGAAGCAGTCTCGGACAGCG
 AGCTCATGGCCAGCAGCACCACGCCGCTCTGCCGGCTGGCTTCTGCCCGGGGCTGCCCGCCCTCG
 CTACCTCTTCCAGAGAAGGTCCAAGCTATGGGGGACCCCGTGGAGAGCCGGGGCTTCCCTGGGGCTGAA
 GACGACAAACCACTGTGATCAGTGAGCTCAGCTCCCCGCTGCAGCAGCTGAACAAGGACACCGCTTCCC
 TGGGGGAGGAACCAAGTTGGTGGCCTGGGCAGCCTGCTGGACCTGCCAAGAAGTGCCTCATCGCAGCAGC
 TCGGCTCTTTCAGCAGCCTCGGTGAGCTGAGCTCCATTTTCAGCGCAGCGCAGCCCCGGGGGCGCGGGC
 GGGGCTCTGTACTCGGTGAGGCCAGTGGCCGCTACCCCGTGGCGAGACGCGCCCCGAGCCCGGTGAAGC

CCGCGTCGCTGGAGCGGGTGGAGGGGCTGGGGCGGGCGCGGGGGCGCAGGGCGGCCCTTCGGCCTCAC
 GCCCCCACCATCCTCAAGTCGTCCAGCTCTCCATCCCGCACGAGCCCAAGGAGGTGCGCTTCGTGGTG
 CGCAGCGTGAGCGCGCGCAGTCGCTCCCCCTCGCCGTCGCCGCTGCCCTCGCCCGCGTCCGGCCCCGGCC
 CCGGCGCCCCCGGCCACGCCGACCTTCCAGCAGAAGCCGCTGCAGCTCTGGAGCAAGTTCGACGTGGG
 CGACTGGCTGGAGAGCATCCACCTAGGCGAGCACCGCGACCGCTTCGAGGACCATGAGATAGAAGGCGCG
 CACCTACCCGCGCTTACCAAGGACGACTTCGTGGAGCTGGGCGTCACGCGCGTGGGCCACCGCATGAACA
 TCGAGCGCGCGCTCAGGCAGCTGGACGGCAGCTGA

>gi|15451914|ref|NP_003924.2| BAI1-associated protein 3; BAI-associated protein 3 [Homo sapiens]

MRPRGAFAAGPPGDLHLGTAIGFAGAIWRSRSPAMSTLLDIKSSVLRQVQVCPFRRRTEQDPGSASAD
 PQEPATGAWKPGDGEFFAHMRLMLKKGEGRQGLPCLEVPLRSGSPAPPEPVDPSLGLRALAPEEVEMLY
 EEALYTVLYRAGTMGPDQVDDEEALLSYLQQVFGTSLEEHTEAIERVRKAKAPTYALKVSVMRANKLLAK
 DPNGFSDPYCMLGILPASDATREPRAQKEQRFGRKGSKRGGPLPAKCIQVTEVKSSTLNPNVWKEHFLFE
 IEDVSTDQLHLDIWDHDDVSLVEACRKLNEVIGLKGMRIFYKQIVKSARANGTAGPTEDHTDDFLGCLN
 IPVREVPVAGVDRWFKLEPRSSASRVQGHCHLVKLITTRQDTAMSQRGRSGFLSHLLLLSHLLRLEHSA
 EEPNSSSWRGELSTPAATILCLHGAQSNLSPLQLAVLHWQVSSRHHQTCTLDYSYLLGLLEDMAHWEAA
 PSLPQEQEESLADSLSAFSEFGLQLLRQLRDYFPATNSTAVHRELELLKCLGKLQLFQPSFEICPFESL
 NMDIAAALKRGNREWDRIKNDKSPREQPGPQRLPGLVVLADAVYDDLQFCYSVYASLPHSILNVDVFTL
 TFRQLERLVAEEAWVLTEELSPKMTLEVASGLFELYLTADLQRFWDSIPGRDSRSLALAGIHAPFLPAV
 KLWFQVLRDQAKWRLQGAVDMDTLEPVDASSRHSSAATAGLCLSHIQELWVRLAWPDPAQAQGLGTQLG
 QDVCEATLFYTELLRKKVDTQPGAAGEAVSEALCVVLNNVELVRKAAGQALKGLAWPEGATGPEGVLPRP
 LLSCTQALDDDLQREAHVTVAHLTSKMGVDIRKYVQHISLSPDSIQNDEAVAPLMKYLDEKLALLNASLV
 KGNLSRVLEALWELLQAILQALGANRDVSADFYSRFHFTLEALVSFFHAEGQGLPLESLRDGYSYKRLKE
 ELRLHKSTRECIEQFYLDKLRQRTLEQNRFGRLSVRCHYEAAEQRLAVEVLHAADLLPLDANGLSDPFV
 IVELGPPHLFPVRSQRTQVKTRTLHPVYDELFFYVPAEACRRRAACVLFTVMDHDLSTNDFAGEAAL
 GLGGVTGVARPQVGGGARAGQPVTLLHLCPRAQVRSALRRLEGRTSKEAQEFVKKLKELEKCMADP

>gi|27500274|ref|XP_113950.3| similar to Munc13-4 protein [Homo sapiens]

MATLLSHPPQRPFLRQAIRRRRVRDLQDPPQMAPEIQPPSHHFSPEQRALLYEDALYTVLHRLGHP
 EPNHVTEASELLRYLQEAHFVEPEEHQQTQVRVRELEKPIFCLKATVKQAKGILGKDVSGFSDPYCLLGI
 EQGVGVPGGSPGSRHRQKAVVRHTIPEETHRTQVITQTLNPNVWDETFILEFEDITNASFHLDMWDLDTV
 ESVRQKLGETDLHGLRRIFKEARKDKGQDDFLGNVLRQLDLRCREDQWYPLEPRTETYPDRGQCHLQF
 QLIHKRRATSASRSQPSYTVHLHLQLVSHVETQHEAGSTSWDGSLSQAATVFLHATQKDLSDFHQS
 MAQWLAYSRLYQSLFPPSSCLLHPITSIEYQWIQRLKAEQQEELAASFSSLLTYGLSLIRFRSVFPLS
 VSDSPARLQSLRLVLVQMCKMAFGELCPNTAPLPQLVTEALQTGTTEWFHLKQHHQPMVQGIPEAGKA
 LLGLVQDVIGDLHQCRQTWDKIFHNTLKIHLFSMAFREQLVAKRVQDHTTVVGDVVSPEMGESLFLY
 ISLKELCQLRMSSSERDGVLDLNFHRWFQPAIPSWLQKTYNEALARVQRAVQMDLVLPLGELTKHSTSA
 VDLSTCFAQISHTARQLDWPDPPEAFMITVKFVEDTCRLALVYCSLIKARARELSSGQKQDQGAANMLCV
 VVNDMEQLRLVIGKLPAQLAWEALEQRVGAVLEQGQLQNTLHAQLQSALAGLGHEIRTGVRTLAEQLEVG
 IAKHIQKLGVRESVLPEDAILPLMKFLEVELCYMNTNLVQENFSSLLTLLWHTLTVLVEAAASQSSSS
 LASNRLKIALQNLICFHAEGCGLPPKALHTATFQALQRDLQLQAASSRELIRKYFCRIQQQAETTSEE
 LGAVTVKASYRASEQKLRLVELLSASSLLPLDSNGSSDPFVQLTLEPRHEFPELAARETQKHKDLHPLFD
 ETFEFLVPAEPCRKAGACLLLTVDYDTLGADDLEGEAFLPLREVPLSGSEEPGEVPQTRLPLTYPAPN
 GDPILQLLEGRKGDREAQVFVRLRRHRAKQASQHALRPAP

>gi|5031915|ref|NP_005836.1| ATP-binding cassette, sub-family C, member 4; canalicular multispecific organic anion transporter (ABC superfamily) [Homo sapiens]

MLPVYQEVKPNPLQDANICSRVFFWWLNPLFKIGHKRRLEEDDMYSVLPEDRSQHLGEELQGFWDKEVLR
 AENDAQKPSLTRAIIKCYWKSylvLGIFTLIEESAKVIQIFLGKIINYFENYDPMDSVALNTAYAYATV
 LTFCTLILAILHHLFYFHVQCAGMRLRVAMCHMIYRKALRLSNMAMGKTTTGQIVNLLSNDVNKFDQVTV
 FLHFLWAGPLQAIATVALLWMEIGISCLAGMAVLIIILLPLQSCFGKLFSSLSKATFTDARIRTMNEVI
 TGIRIIKMYAWKSFSNLITNLRKKEISKILRSSCLRGMNLASFFSASKIIVFVTFVTVLLGSVITASR
 VFVAVTLYGAVRLTVTLFFPSAIERVSEAIVSIRRIQTFLLLDLDEISQRNRQLPSDGKKMVHVQDFTAFWD

KASETPTLQGLSFTVRPGELLAVVGPVGAGKSSLLSAVLGELAPSHGLVSVHGRIAYVSQQPWVFSGTLR
 SNILFGKKYEKERYEKVAKACALKKDLQLEEDGDLTVIGDRGTTLSGGQKARVNLARAVYQDADIYLLDD
 PLSAVDAEVSRLHFELCICQILHEKITILVTHQLQYLKAASQILILKDGKMQKGTYTEFLKSGIDFGSL
 LKKDNEESEQPPVPGTPTLRNRTFSESSVWSQQSSRPSLKDGALESQDTENVVPTLSEENRSEGVGFQA
 YKNYFRAGAHWIVFIFLILLNTAAQVAYVLQDWWSYWANKQSMNLVTVNGGNNVTEKLDLNLWYLGIIYSG
 LTVATVLFGIARSLLVFYVLVNSSQTLHNKMFESILKAPVLFDRNPIGRILNRFSDKIGHLDDLLPLTF
 LDFIQTLQVVGVSVAVAVIPWIAIPLVPLGIIFIFLRRYFLETSRDVKRLESTTRSPVFSHLSSSLQG
 LWTIRAYKAEERCQELFDAHQDLHSEAWFLFTTSRWFVAVRLDAICAMFVIVAFGSLILAKTLDAQVVG
 LALSVALTLMGMFQWCVRQSAEVENMMISVERVIEYTDLEKEAPWEYQKRPPPAPWPHEGVIIFDNVNFMY
 SPGGPLVLKHLTALIKSQEKVGIVGRTGAGKSSLSALFRLSEPEGKIWDKILTTEIGLHDLRKKMSII
 PQEPVLTGTMRKNLDPFKEHTDEELWNALQEVQLKETIEDLPGKMDTELAESGSNFSVGQRQLVCLARA
 ILRKNQILIIIDEATANVDPRTDELQKKIREKFAHCTVLTIAHRLNTIIDSDKIMVLDSGRLKEYDEPYV
 LLQNKESLFYKMQVQLGKAEAAALTETAKQVYFKRNYPHIGHTDHMTNTSNGQPSTLTIFETAL

>gi|18596951|ref|XP_058661.2| similar to CG11412-PA [Homo sapiens]
 MAEVPFGSSLLPPPAPPAPAAVEPRCPFPAGAALACCSEDEEDDEEHEGGGSRSPAGGESATVAAKGHP
 CLRCPPPPQEQQLNGLISPELRHLRAAASLSKSVLSVAEVAATATPDGGRATATKGAGVHSGERPPH
 SLSSNARTAVPSPVEAAAASDPAAARNGLAEGTEQEEEEDEQVRLSSSLTADCSLRSPSGREVEPGE
 RTIRYVRYESELQMPDIMRLITKDLSEPYSIYTRYFIHNPQLCFLAMVGECEVGAIVCKLDMHKMKMR
 RGYIAMLAVDISKYRRNGIGTNLVKKAIYAMVEGDCDEVVLETEITNKSALKLYENLGFVRDKRLFYYLNL
 GVDALRLKLWLR

>gi|33328294|gb|AAQ09597.1| NS5ATP2 [Homo sapiens]
 MASVSSATFSGHGARSLLQFLRLVGQLKRVPRGTGWVYRNVQRPESVSDHMYRMAVMAMVIKDDRNLNKDRC
 VRLALVHDMAECIVGDIAPADNIPKEEKHREEEAMKQITQLLPEDLRKELYELWEEYETQSSAEAKFVK
 QLDQCEMILQASEYEDLEHKPGRQLQDFYDSTAGKFNHPEIVQLVSELEAERSTNIAAAASEPHS

>gi|22202624|ref|NP_665877.1| glutathione transferase zeta 1 isoform 1;
 glutathione s-transferase Zeta 1; maleylacetone isomerase; glutathione
 S-alkyltransferase; glutathione S-aryltransferase; S-
 (hydroxyalkyl)glutathione lyase; glutathione S-aralkyltransferase;
 maleylacetate isomerase [Homo sapiens]
 MQAGKPILYSYFRSSCSWRVRIALALKGIDYETVPINLIKDGQQFSKDFQALNPMKQVPTLKIDGITIH
 QSLAIIEYLEEMRPTPRLLPQDPKKRASVRMISDLIAGGIQPLQNLVSVLKQVGEEMQLTWAQNAITCGFN
 ALEQILQSTAGIYCVGDEVMTADLCLVPQVANAERFKVDLTPYPTISSINKRLLVLEAFQVSHPCRQPD
 PTELRA

>gi|21361433|ref|NP_055079.2| serine/threonine kinase 18; Snk Akin
 Kinase [Homo sapiens]
 MATCIGEKIEDFKVGNLLGKGSFAGVYRAESIHSGLVAIKMIDKKAMYKAGMVQRVQNEVKIHCQLKHP
 SILELYNYFEDSNYVYLVLEMCHNGEMNRYLKNRVKPFSENEARHFMHQIITGMLYLHSHGILHRDLTSL
 NLLLTRNMNIKIADFGLATQLKMPHEKHYTELCGTPNYISPEIATRSAGHLESDVWSLGCMTYLLIGRPP
 FDTDTVKNTLKNKVLADYEMPTFLSIEAKDLIHQLLRNPADRLSLSSVLDHFPMSRNSSTKSKDLGTVE
 DSIDSGHATISTAITASSSTSISGSLFDKRRLLIGQPLPNKMTVPKPKNSSTDFSSSGDGNSTFYQWGNQ
 ETSNSGRGRVIQDAEERPHSRYLRRAYSSDRSGTSNRQSQAKTYTMERCHSAEMLSVSKRSGGGENEERY
 SPTDNNANIFNFFKEKTSSSSGSFERPDNNQALS NHLCPGKTPFPFADPTPQTETVQQWFGNLQINAHLR
 KTTEYDSISPNRDFQGHDPDLQKDTSKNAWTDTKVKKNSDASDNAHSVQQNTMKYMTALHSKPEIQQEC
 VFGSDPLSEQSKTRGMEFPWGYQNRTLRSITSLVAHRLKPIRQKTKKAVVSILDSEEVCELVKEYASQ
 EYVKEVLQISSDGNITITIIYPNGGRGFPLADRPSPPTDNI SRYSDNLPEKYWRKYQYASRFVQLLRSSK
 PKITYFTRYAKCILMENSFGADFEVWFYDGVKIHKTEDFIQVIEKTGKSYTLKSESEVNSLKEEIKMYMD
 HANEHGRICLALESIISEEEKTRSAFPFPIIIGRKPGSTSSPKALSPPPSVDSNYPTRDRASFNRMMH
 SDASPTQAPILNPSMVTNEGLGLTTTASGTDISSNSLKDCLPKSAQLLKS FVKNVGWATQLTSGAVVWQ
 FNDGSQVLVQAGVSSISYTS PNGQTRYGENEKLDPYIKQLQLCLSSILLMFSNPTPNFH

>gi|20555151|ref|XP_166453.1| similar to KIAA1855 protein [Homo
 sapiens]
 MDLLTRENVALKVESAAQPKQVLKMEVAVLKKLQGGKDHVCRFIGCGRNEKFNYYVVMQLQGRNLADLRRSQ

PRGTFTLSTTLRLGKQILESIEAIHSGVGLHRDIKPSNFAMGRLPSTYRKCYMLDFGLARQYTNTTGDVR
 PPRNVAGFRGTVRYASVNAHKNRHDDLWSLFYMLVEFAVGQLPWRKIKDKEQVGMKEKEYEHRMLL
 KHMPSEFHLFLDHIASLDYFTKPDYQLIMSVFENSMKERGIAENAFDWEKAGTDALLSTSTSTPPQONT
 RQTAAMFGVVNTPVPGDLLRENTEDVLQGEHLSDQENAPPILPGRPSEGLGPSPHLVPHPGGPEAEVWE
 ETDVNRNKLRLINIGKSPCVEEEQSRGMGVPSSPVRAPPDSTPTPVRSLRYRRVNSPESERLSTADGRVEL
 PERRSRMDLPGSPSRQACSSQPAQMLSVDTGHADRQASGRMDVSASVEQEALSNAFRSVPLAEEDFDISK
 EWVIDKETELKDFPPGAEPSTSGTTDEEPEELRPLPEEGEERRRLGAEPTRPRGRSMQALAEEDLQHL
 PPQPLPPQLSQGDGRSETSPPTPGSPSHSPLHSGPRPRRRESPTGPRQVFSVAPPFEVNGLPRAVPL
 SLPYQDFKRDLSDYRERARLLNRVRRVGFSHMLLTPQVPLAPVQPQANGKEEEEEEEEEEEEEEEEE
 EEEEEEEEEEEEEEEEEEEEEAAAAVALGEVLGPRSGSSSEGSERSTDRSQEGAPSTLLADDQKESRGRAS
 MADGDLEPEEGSKTLVLVSPGDMKSPVTAELAPDPLGTLAALTPQHERPQPTGSQLDVSEPGTLSSVL
 KSEPKPPGPGAGLGAGTVTTGVGGVAVTSSPFTKVERTFVHIAEKTHLNMSSGGQALRSEEFSGAGELG
 LELASDGGAVEEGARAPLENGLALSGLNGAEIEGSALSGAPRETPEMATNSLPNGPALADGPAPVSPLE
 PSPEKVATISPRRHAMPGRSRPRSRI PVLLSEEDTGSEPSGSLSAKERWSKRARPQQDLARLVMEKRQGRIL
 LLRLASGASSSSSEEQRRASETLSTGTGSEEDTPASEPAAALPRKSGRAAATRSRI PRPIGLRMPMPVAAQ
 QPASRSHGAAPALDTAITSRLQLQTPPGSATAADLRPKQPPGRGLGPGRAQAGARPPAPRSPRLPASTSA
 ARNASASPRSQSLSRRESPPSHQARPGVPPPRGVPPARAQPDGTPSPGGSKKGRGKLQAQRATTKGRA
 GGAEGRAGAR

>gi|28466991|ref|NP_775771.2| tau tubulin kinase 2; tau-tubulin kinase
 [Homo sapiens]

MESVEKDWGWLWRNLRLCLGHAHQKCCTEGGISSTTKTSSENGSCCFEKAARKDHVCRFIGCGRNDRFN
 YVVMQLQGRNLADLRSSQSRGTFTISTTLRLGRQILESIESIHSVGFLHRDIKPSNFAMGRFPSTCRKCY
 MLDFGLARQFTNSCGDVRPPRAVAGFRGTVRYASINAHNRHDDLWSLFYMLVEFVVGQLPWRKIK
 DKEQVGSIKERYDHRMLMLKHLPPFEFSIFLDHISSLDYFTKPDYQMSIRVTRKSYKVSTSGPQAFNSLSYT
 SGPGACISSSSFSRMGSSSFRGGLGAGYGASGGITTITVNQSLLSPLNLEVDPNIAVVRTQEEKQIKTL
 NNKFFSFIDKVRFLQEQNKMLETKWSLVQQQKMARSNMDNMFESYINNLKWQLETGLQEKLEAEELGNM
 HGLVEDFRNKYEVEISKCTEMENEFVLKEYVDEAYNMKMELESSLKELTAKISFLRQLYEEEIGSCSPR
 SRIHLWCCSWTTAASWTWTASSLRSPGFPGGCHADTEQRGELAIKDANAKLSELEAALQLASQDMARQ
 LREYQELMNVKLALDIKIATYRKLEGEESWLESGMQNMSIHMKTSSYAGGQSLAYGGLTSPGLSYGLG
 SSFGSGMGSSSFSHSTSSRAVVMKKIETRDGKLVSSESNVLPNLRITATLLTSVFDNSIKTFGVIESDPF
 DWEKTGNDGSLTTTTSTTPQLHTRLTPAAIGIANATPIPGDLLRENTDEVFPDEQLSDGENGIPVGVSP
 DKLPGLSLGHRPQEKDVWEEMDANKNKIKLGICKAATEEENSHGQANGLLNAPSLGSPIRVRSEITQPD
 DIPLVRKLRSIHSFELEKRLTLEPKPDTDKFLETCELMQKDTTSAGKESILPALLHKPCVPAVSRTDHIW
 HYDEEYLPDASKPASANTPEQADGGGNGFIAVNLSSCKQEIDSKEWVIVDKEQDLQDFRTNEAVGHKTT
 GSPSDEEPEVLQVLEASPQDEKLQLGPAENDHLKKETSGVVLALSAEGPPTAASEQYTDRLQLPGAAS
 QPIAATPTSLMEAQAEGLPTAITI PRPSVASTQSTSGSFHCGQPEKKDLQPMPTVELYSPRENFSGLV
 VTEGEPSPSGSRDTLGLQIDHIGDMPLNIRESNKSQDLGPKELPDHNLVVRREFENLPGETEEKSILLE
 SDNEDEKLSRGQHCIEISSLPGLVIVEKDHSATTEPLDVTKTQTFSVVNPQDKNNEIMKLLTVGTSEIS
 SRDIDPHVEGQIGQVAEMQKNKISKDDIMSEDLPGHQGDLSTFLHQEGKREKITPRNGELFHCVSENEH
 GAPTRKDMVRSSFVTRHSRI PVLAQEIDSTLESSSPVSAKEKLLQKKAYQPDVLKLLVEKRQKFSFLGDL
 SSASDKLLEEKLATVPAPFCEEEVLTPFSRLTVDSHLSRSAEDSFLSPIISQSRKSKI PRPVSWVNTDQV
 NSSTSSQFFPRPPPGKPPTRPGVEARLRRYKVLGSSNSDSFLSRLAQILQNGSQKPRSTTQCKSPGSPH
 NPKTPPKSPVPRRSPSASPRSSSLPRTSSSSPSRAGRPHHDQRSSSPHLGRSKSPSPSHSGSSSSRRSCQ
 QECHKPSKNGLKSGSLHHHSASTKTPQGKSKPASKLSR

>gi|4505467|ref|NP_002517.1| 5' nucleotidase, ecto; Purine 5-Prime-
 Nucleotidase; 5' nucleotidase (CD73); ecto-5'-nucleotidase [Homo
 sapiens]

MCPRARAPATLLALGAVLWPAAGAWELTILHTNDVHSRLQTSSESSKCVNASRCMGGVARLFTKVQQ
 IRRAPENVLLLDAGDQYQGTIWFTVYKGAEVAHFMMNALRYDAMALGNHEFDNGVEGLIEPLLEAKFPIL
 SANIKAKGPLASQISGLYLPYKVLPGVDEVVGIVGYTSKETPFLSNPGTNLVFEDEITALQFEVDKLT
 NVNKIIALHSGFEMDKLIAQKVRGVDDVVGGHSNTFLYTGNPPSKEVPAGKYFFIVTSDDRKVPVQA
 YAFGKILGYLKI EFDERGNVISSHGNPILLNSSIPEDPSIKADINKWKIKLDNYSTQELGKTIYVLDGSS
 QSCRFRECNMGNLICDAMINNNLRHTDEMFWNHVSMCILNGGGIRSPIDERNNGTITWENLA AVL PFGGT
 FDLVLQKGSTLKKAFESVHRYGQSTGEFLQVGGIHVVYDL SRKPGDRVVKLDVLC TKCRVPSYDPLKMD
 EVYKVILPNFLANGGDGFQMIKDELLRHDSDGDQDINNVSTYISKMKVIYPAVEGRIFSTGSHCHGSFSL

EX03-092P

IFLSLWAVIFVLYQ

>gi|17402896|ref|NP_478123.1| RAD51 homolog C isoform 1; DNA repair protein RAD51 homolog 3; yeast RAD51 homolog 3 [Homo sapiens]
MRGKTFRFEMQRDVLSFPLSPAVRVKLVSAAGFQTAEELELVKPSSELSKEVGISKAELETQIIRRECLT
NKPRYAGTSESHKCTALELLEQEHTQGFIIITFCSALDDILGGGVPLMKTTEICGAPGVGKTQLCMQLAV
DVQIPECFGGVAGEAVFIDTEGSFMDRVVDLATAACIQHLQLIAEKHKGEEHRKALEDFTLDNLSHIYY
FRCRDYTELLAQVYLLPDLSEHSKVRVIVDGIAPFRHDLDDLRLTRLLNGLAQQMISLANNHRLAV
ILTNQMTTKIDRNQALLVPALGESWGHAATIRLIFHWDRKQRLATLYKSPSQKECTVLFQIKPQGFRDTV
VTSACSLQTEGSLSTRKRSRDPEEL

>gi|4885657|ref|NP_005422.1| X-ray repair cross complementing protein 2; X-ray repair, complementing defective, repair in Chinese hamster; DNA repair protein XRCC2 [Homo sapiens]
MCSAFHRAESGTELLARLEGRSSLKEIEPNLFADEDSPVHGDILEFHGPEGTGKTEMLYHLTARCILPKS
EGGLEVEVLFIDTDYHFDMLRLVTILEHRLSQSSEEIIKYCLGRFFLVYCSSSTHLLLTLYSLESMFCSH
PSLCLLILDLSAFYWDVNGGESVNLQESTLRKCSQCLEKLVNDYRLVLFATTQTIMQKASSSSEEPS
HASRRLCDVDIDYRPYLCKAWQQLVKHRMFFFSKQDDSQSSNQFSLVSRCLKSNSLKKHFFIIGESGVEFC

>gi|11968152|ref|NP_057232.1| SH3 and multiple ankyrin repeat domains 1; somatostatin receptor-interacting protein [Homo sapiens]
MTHSPATSEDEERHSASECEGGSESDDSSPDGPRGRGTRGQSGAGPSLASVRGLQGRSMSGVPPDAHF
SMMVFRIGIPDLHQTKCLRFNPDAIWTAKQVLCALSESLQDVLNYGLFQPATSGRDANFLEERLLRE
YPQSFEKGVYPYLEFRYKTRVYKQTNLDEKQLAKLHTKTGLKKFLEYVQLGTSDKVARLLDKGLDPNYHDS
DSGETPLTLAAQTEGSVEVIRTLCLGGAHIDFRARDGMTALHKAACARHCLALTALDLGGSPPNYKDRRG
LTPLFHTAMVGGDPRCELLLFNRAQLGIADENGWQEIHQACQRGHSQHLEHLLFYGAEPGAQNASGNTA
LHICALYNKETCARILLYRGADKDVKNNGQTPFQVAVIAGNFELGELIRNHREQDVVPFQESPKYAARR
RGPPGTGLTVPPALLRANS DTSMALPDWMVFSAPGAASSGAPGPTSGSGGQSQPSAPTTLSSGTLRSAS
SPRGARARSPSRGRHPEDAKRQPRGRPSSSGTPREGPAGGTGGSGGPGGSLGSRGRRRLYSAVPGRSFM
AVKSYQAQAEGEISLSKGEKIKVLSIGEGGFWEQVKGVRVGFPSDCLEEVANRSQESKQESRSDAKRL
FRHYTVGSGYDSFDAPSLMDGIGPGSDYIIKEKTVLLQKKDSEGFVLRGAKAQTPIEEFTPTPAFPALQ
YLESVDEGGVAWRAGLRLMGDFLIEVNGQNVVKVGRVQVNMIRQGGNTLMVKVVMVTRHPDMDEAVHKKA
PQQAQRLPPPTISLRSKSMTSELEEMEYEQQPAPVPSMEKKRTVYQMALNKLDEILAAAQQTISASESPG
PGGLASLGKHRPKGFFATESFDPHRAQPSYERPSFLPPGPGMLRLQKSIGAAEDDRPYLAPPAMKFSR
SLSVPGSEDIPPPPTTYPPEPPYSTPTVPSSSGLRTPSPRGPFNPGSGGPLPASSPATFDGSPFPDTRV
GSREKSLYHSGPLPPAHHPHPPHHHHHAPPQPHHHHAHPHPPEMETGGSPDDPPRLALGPQPSLRGW
RGGGPSPTPGAPSPSHHGSAGGGGSSQGPALRYIQLPPRAASAAMYVPARSGRGRKGPLVKQTKVEGEP
QKGGGLPPAPSPSSPASQPPPAVAAPSEKNSIPIPTIIKAPSTSSSGRSSQSSSTEAEPTTQPEPTGG
GGGGSSPSPAPAMSPVPPSPSPVPTPASPSGPATLDFTSQFGAALVGAARREGGWQNEARRRSTLFLST
DAGDEDGGDGLGTGAAPGPRLRHKSIDEGMFAEPYLRLSAGSGAGYGGYGAGSRAYGGGGSSAFT
SFLPPRPLVHPLTGKALDPASPLGLALAAREALKESSSEGGAPQPPRPPSPRYEAPPPTPHHSPHAP
HEPVLRLWGAASPPDPARRELGYRAGLGSQESLPASPPAARRSLHRLPPTAPGVGPLLQLGTEPPAPH
PGVSKPWRSAAPEEPERLPLHVRFLENCQPRAPVTSGRGPPSEDGPGVPPPSPRRSVPPSPSTSPRASEN
GLPLLVLPPPAPSDVDEGEFLFVEPLPPPLEFSNSFEKPESPLTPGPPHPLPDTAPATPLPPVPPPAV
AAPPTLDSTASSLTSDSEVATLTQGASAAPGDHPHPPGPPAPAPAPAPQPGPDPPPGTDSGIEEVDS
RSSSDHPLETISASTLSSLSAEGGGSAGGGGAGAGVASGPPELLDTYVAYLDGQAFGGSSTPGPPYPPQ
LMTPSKLRGRALGASGLRPGPSGGLRDPVTPTSTVSVTGAGTDGLLALRACSGPPTAGVAGGPVAVEP
EVPVPLPTASCLPRKLLPWEEGPGPPPPPLPGPLAQPPASALATVKASIISELSSKLQQFGGSSAAGGA
LPWARGGSGGGGSDSHGGASYVPERTSSLQRQLSDSSQSLLSKPVSSSLFQNWPKPPLPLPTGTGVSP
TAAAPGATSPSASSSSTSTRHLQGVFEMRPPLLRAPSPSLLPASEHKVSPAPRPSLPILPSGPLYP
GLFDIRGSPTGGAGGSADPFAPVFVPPHPGISGGLGGALSGASRSLSPTRLLSLPPDKPFGAKPLGFWTK
FDVADWLEWLGLAEHRAQFLDHEIDGSHLPALTKEYVDLGVTRVGHMNMIDRALKFFLER

>gi|19743794|ref|NP_036441.1| SH3 and multiple ankyrin repeat domains 2 isoform 1; cortactin SH3 domain-binding protein; cortactin binding protein 1; GKAP/SAPAP interacting protein [Homo sapiens]
MMNVPGGAAAVMMTGYNNGRCPRNSLYSDCIIEEKTVVLLQKKDNEGFVLRGAKADTPIEEFTPTPA

FPALQYLESVDEGGVAWQAGLRTGDFLIEVNNENVVKVGHQVVMIRQGGNHLVLKVTVTRNLDPDDT
 ARKKAPPPPKRAPTTALTLRSKSMTELEELVDKDKPEEIVPASKPSRAAENMAVEPRVATIKQRPSSRC
 FPAGSDMNSVYERQGIAMVTPTPVPGSPKAPFLGIPRGTMRRQKSIDSRIFLSGITEEERQFLAPPMLKFT
 RSLSPMDTSEDIPPPQSVPPSPPPPSPTTYNCPKSPTPRVYGTIKPAFNQNSAAKVSPATRSDTVATMM
 REKGMFYRRELDRLSLDSEDLYSRNAGPQANFRNKRQMPENPYSEVGKIASKAVYVPAKPARRKGMVLK
 QSNVEDSPEKTCISIPITIIIVKEPSTSSSGKSSQSSMEIDPQAEPPPSQLRPDESLTVSSPFAAAIAGA
 VRDREKRLARRNSPAFLSTDLGDEDVGLGPPAPRTRPSMFEEGDFADESDAEQLSSPMPSATPREPEN
 HFVGGAESAAPGEAGRPLNSTSKAQGPESPAVPSASSGTAGPGNYVHPLTGRLLDPSSPLALALSARDR
 AMKESQOGPKGEAPKADLNKPLYIDTKMRPSLDAGFPPTVTRQNTRGPLRRQETENKYETDLGRDRKGDDK
 KNMLIDIMDTSQQKSAGLLMVHTVDATKLDNALQEEDEKAEVEMKPDSSPSEVPEGVSETEGALQISAAP
 EPTTVPGRTIVAVGSMEEAVILPFRIPPPPLASVDLDEDFIFTEPLPPPLEFANSFDIPDDRAASVPALS
 DLVKQKKSPTPQSPSLNSSQPTNSADSKKPASLSNCLPASFLPPPEFSDAVADSGIEEVDSRSSSDHLE
 TTSTISTVSSISTLSSEGGENVDCTVYADGQAFMVDKPPVPPKPKMKPIIHKSNALYQDALVEEDVDSF
 VIPPPAPPPPPGSAQPGMAKVLQPRTSKLWGDVTEIKSPILSGPKANVISELNSILQQMNRKLAKEGEG
 LDSPMGAKSASLAPRSPEIMSTISGTRSTTVTFTVRPGTSQPIITLQSRPPDYESTSGTRRAPSPVVSPT
 EMNKETLPAPLSAATASPSPALSDVFLSPSQPPSGDLFGLNPAGRSRSPSPSILQQPISNKPFTTKPVHL
 WTKPDVADWLESNLNGEHKEAFMDNEIDGSHLPNLQKEDLIDLGVTRVGHRMNIERALKQLDDR

>gi|37564792|ref|XP_037493.2| similar to Proline rich synapse
 associated protein 2 [Homo sapiens]

MQLSRAAAAAAFAEPPEPLSPAPAPAPPPGPLPRSAADGAPAGGKGGPGRARRRRGPGKPGPWGGAS
 AAKCLRLDPAAPVWAAKQRLCALNHSIQDALNYGLFQPPSRGRAGKFLDEERLLQEYPPNLDTPLPYLE
 FRYKRRVYAQNLIIDDKQFAKLHTKANLKKFMDYVQLHSTDKVARLLDKGLDPNFHDPDSGECPLSLAAQL
 DNATDLLKVLKNGGAHLDFRTRDGLTAVHCATRQNAALTTLLDLGASPDYKDSRGLTPLYHSALGGGD
 ALCCLELLHDHAQLGITDENGWQEIHQACRFHGVHLEHLLFYGADMGAQNASGNTALHICALYNQESCA
 RVLLFRGANRDVRNYSQTAFQVAIIAGNFELAEVIKTHKDSDVVPFRETPSYAKRRRLAGPSGLASPRP
 LQRSASDINLKGEAQPAASPGPSLRSLPHQLLLQRLQEEKDRDRDADQESNISGPLAGRAGQSKISAVPG
 RKFIKAVKAHSPQVLSIGEGGFWEFTVKGRGTGWFADCVVEEQMRQHDTRPETREDRTKRLFRHYTVGSYD
 SLTSHSDYVIDDKVAVLQKRDHEGFGFVLRGAKAETPIEEFTPTPAFPALQYLESVDVEGVAWRAGLRTG
 DFLIEVNGVNVVKVGHKQVVALIRQGGNRLVMKVSVTRKPEEDGARRRRAMKKFASSRNLNKILAQCDS
 SREYEEIQAVERKWLHLATPRRLDDRASKASLFFAAPPPKRAPSTTLTLRSKSMTELEELEKLDDEM
 LAAAAEPTLRPDIAADDSRAATVKQRPTSRRITPAEISSLFERQGLPGPEKLPGLSRKGIPTKSVGEDE
 KLASLLEGRFPRSTSMQDPVREGRGIPPPQTAAPPPPPAPYFDSGPPPAFSPPPPPGRAYDTRVSSFKP
 GLEARLGAGAAGLYEPGAALGPLPYPERQKRARSMIILQDSAPESGDAPRPPPAATPPERPKRRRPPPGP
 DSPYANLGAFSASLFAPSKPQRRKSPLVKQLQVEDAQERAALAVGSPGPGGGSFAREPSPTHRGPRPGGL
 DYAGDGPGLAFGGPGPAKDRRLEERRRSTVFLSVGAIEGSAPGADLPQLPSRSIDERLLGTGPTAGRD
 LLLPSPVSALKPLVSGPSLPGSGSTFIHPLTGKPLDPSSPLALALAAERERASQAPSRSPTPVHSPDAD
 RGPGLFVDVQARDPERGSLASPAFSPRSPAWIPVPARREAEKVPREERKSPEDKKSMLSVLDTSLQRP
 GLIVVHATSNGQEP SRLGGAEERPGTPELAPAPMQSAVAEPLPSAPRAQPPGGTPADAGPGQGSSEEP
 ELVFAVNLPPAQLSSSDEETREELARIGLVPPPEEFANGVLLATPLAGPGSPPTTVPSPASGKPSSEPP
 APESAADSGVEEADTRSSSDPHLETTSTISTVSSMSTLSSSESGELTDHTSFADGHTFLEKPPVPPKPK
 LKSPGLGKGPVTRDPLLKQSSDSELMAQQHHAASAGLASAAGPARPRYLQRRSKLWGDVPESRGLPGPE
 DDKPTVISELSSRLQQLNKDTRSLGEEPVGGLGSLDDPAKKSPIAAARLFSSLGELSSIQAQRSPGPGG
 GASYSVRPSGRYPVARRAPSPVKPASLERVEGLGAGAGGAGRPFGLTPTPTILKSSSLSI PHEPKEVRFV
 RSVSARSRSPSPSLPSPASGPGGAPGPRRPFQKPLQLWSKFDVGDWLES IHLGEHRDRFEDHEIEGA
 HLPALTKDDFVELGVTRVGHRMNIERALKQLDGS

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/039549

International filing date: 23 November 2004 (23.11.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/524,587
Filing date: 24 November 2003 (24.11.2003)

Date of receipt at the International Bureau: 10 January 2005 (10.01.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse